

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 05:48:05 ; Search time 2447.42 Seconds
(without alignments)
6441.250 Million cell updates/sec

Title: US-09-245-198a-1
Perfect score: 1168
Sequence: 1 ggtgctgagcctggcctg.....ataatcatgtctcttc 1168

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST :
1: em_estba :
2: em_esthum :
3: em_estin :
4: em_estnu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_hic :
9: gb_estl :
10: gb_est2 :
11: gb_hic :
12: gb_gss :
13: em_gss_hum :
14: em_gss_inv :
15: em_gss_pln :
16: em_gss_vrl :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	883.8	75.7	1033	11	AK020909	Mus muscu
2	625.8	53.6	918	10	BE577781	602092080
3	519.4	44.5	731	10	BI871711	603395825
4	507.4	43.4	728	10	BI870393	603395641
5	504	43.2	561	9	AW763237	ur70609..r
6	497	42.6	533	10	BE628951	un30c03.y
7	489.6	41.9	650	10	BE628951	602420160
8	481.2	41.2	584	9	AW917574	EST348878
9	480.4	41.1	777	10	BI819200	603034614
10	473	40.5	834	10	BI766766	603056866
11	456.4	39.1	828	10	BI596681	603243254
12	437.8	37.5	471	9	AA221610	my18609..r
13	390.6	33.4	445	9	AA870722	vq25g07..r
14	360	30.8	360	10	BE654876	UI-M-BH0-
15	353	30.2	405	9	AI854476	UI-M-BH0-
16	331	28.3	367	10	BI111534	602895883
17	325.4	27.9	376	9	AA792068	vm69d09..r

18	309	26.5	474	10	BI965174	ld34e07.y
19	306	26.2	315	10	BF466521	UI-M-CG0P
20	299.4	25.6	538	10	BF821434	MR1-R7003
21	297.2	25.4	400	9	AI152313	u087h02..r
22	292.8	25.1	493	10	BE307031	601087888
23	289.6	24.8	433	10	BF283688	EST448279
24	289.4	24.8	542	10	BF041509	BP250025B
25	285.2	24.4	443	10	BC378802	UI-R-CV1-
26	284.4	24.3	785	10	BI762980	603047966
27	283.8	24.3	404	10	BE044430	BP250013A
28	278.4	23.8	897	10	BI730298	603350276
29	274.4	23.5	412	9	AI010416	EST204867
30	271	23.2	581	10	BI738634	603358846
31	270.4	23.2	402	10	BF410871	UI-R-CN0-
32	270.2	23.1	292	9	BB268794	BB268794
33	258.2	22.1	558	10	BM484863	53865 MA
34	255.2	21.8	394	10	BG376757	UI-R-CU0-
35	243.6	20.9	371	10	BI300370	UI-R-CV2-
36	242.4	20.8	441	10	BI967060	UI-R-CV2-
37	242.2	20.7	374	9	AA800970	EST190467
38	240	20.5	240	9	AW764050	ur70409..x
39	237	20.3	894	10	BI908274	603068526
40	236.6	20.3	342	9	AA637970	vr30a10..r
41	232.4	19.9	322	10	BI299581	UI-R-CV2-
42	228.8	19.6	318	10	BI300642	UI-R-CV2-
43	227.2	19.5	345	9	AA875031	UI-R-E0-C
44	227.2	19.5	345	10	R55379	yj77a08..r1
45	226.4	19.4	320	9	AI408018	EST236308

ALIGNMENTS

RESULT 1	AK020909	1033 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK020909				
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length cDNA library, member 12, full insert sequence.				
ACCESSION	AK020909				
VERSION	AK020909.1	GI:12861640			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult retina cDNA to mRNA, clone:lib-RIKEN full-length enriched mouse cDNA library clone:AS30030D13.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2 (sites)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	2049374				
PUBMED	11042159				
REFERENCE	3 (sites)				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachi, K., Yoneda, Y., Ishikawa, T., Ozawa, Y., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				

MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 1033)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baladrelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawaji,K., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Ouakenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schiraldi,L., Shibata,K., Shibata,Y., Shinaigawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-AUG-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGATCCAGACCTCTTTTTCCTTTTTCCTTA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGACTTCTCAGTTAATTAAATTATGCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Guslinch (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Retina RNA was provided by Stefano Guslinch (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.
FEATURES	source Location/Qualifiers 1..1033 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1911867" /db_xref="taxon:10090" /clone="A930030DI3" /tissue_type="retina" /clone_id="RIKEN Full-length enriched mouse cDNA library" /dev_stage="adult"
gene	1..433 /gene="tnfrsf12" <1..433 /gene="tnfrsf12" /note="data source:MGD, source key:MGI:1196259, evidence:ISS putative tumor necrosis factor (ligand) superfamily, member 12"
CDS	/codon_start=2 /protein_id="BAB32249.1"

Query Match	75.7%	Score 883.8	DB 11	Length 1033
Best Local Similarity	99.4%	Pred. No. 6.9e-205		
Matches 918	Conservative 0	Mismatches 3	Indels 3	Gaps 3
QY 248	atcgagcccccctctctgaggttcacatccctcggccagagacaggaatggagcaacagcagtg	307		
Db 2	ATTGAGAGCCCATTTATGAGGTTTCATCCTCGGCGAGAGAGATGGAGCACAAGAGTG	61		
QY 308	gatggagacatgtagtgctgagggaagagacaacaatcaacagctccagccctctg	367		
Db 62	GATGGAGACAGTAGAGGCTGGGGAAGAGACCAAAATCAACACCTCCAGCCCTTGGGCTAC	121		
QY 368	gaccgcagagatgggggaattcaacagtcacagagcctgagcctacacctgactg	427		
Db 122	GACCCCGCATTTGGGGAATTTACACTCATTCAGAGGCTGGGCTCTACTACTGTACTGTAG	181		
QY 428	gtgcaatttgatggaggaagagctgtctacatgaagctggaactgtggtgaaagtg	487		
Db 182	GTGCATTTGATGAGGGAAGAGGCTGTCACTCCATGAAGCTGCTGCTGTAACGGTGTG	241		
QY 488	ctggccctgcgctgccttggaagaattctccacagacagagaagaaagcttcctgggcccag	547		
Db 242	CTGGCCCTGCCTGCTCTGGAAGAAATCTCAGCCACAGACAGACTCTCTGGGCCCCAG	301		
QY 548	ctccgcttgatggcagggtgctctggagctgtctgcgcgtcgagcagaggtctctccctcgatc	607		
Db 302	CTCCGTTTGTGCAAGGTGTCTGGGCTGTGGCGGTGGCGGAGGGATCTTCCCTTCGGATC	361		
QY 608	cgcaaccctccctgggctcattcttaagctctgccccctctcaactactttggactctt	667		
Db 362	CGCACCCCTCCCTGGGCTCATCTTTAAAGCTCGCCCTTCCTCAACTACTTTGGACTCTTT	421		
QY 668	caagctcaactgagggcctgtctccacagattctcttaactcttccctggctccagagagc	727		
Db 422	CAAGTTCACTAGGGGCTTGCTGTCCAGATTCCTTTAACTTTCCTGGGCTCCAGAGAC	481		
QY 728	atcacacacactccctaccccaaccccaactccctccacccctcgctgctctctgtgtccag	787		
Db 482	ATCACACACCTTCTTACCTCCACCCCACTCTCCACCCCTTACCTGCTCTCTTGGTCCAG	541		
QY 788	ctcctt-ctctccctaaagcagcagagctgtgtcacatg-cttccatctccacagagct	845		
Db 542	TCTGTCTCTCTCTCAAAAGGACGACGAGCTTGTTCACATGTTTTCATTCCACAGAGCT	601		
QY 846	atcctgtgctctc-ctaaatcccatccacacacaaactatccaaactcagctccaaa	904		
Db 602	ATCCTTGTCTTCTTTAACTATCCATCCACACCACTATCCACTCTCACTGACTGCCAAA	661		
QY 905	ggccctacttatcccttgctcccccacacacacacagcagcaagctgttatgactgtg	964		
Db 662	GCCTCTACTTATCTCTGACTCCCCACCCACTCACCCACACAGCTGTTATGTAGCTTTGT	721		
QY 965	gaaccagagcactgagatgagctcgacacctgtgacaggaagccagagaacactggagctagg	1024		
Db 722	GCACGACGACATGAGATGGGCTGGACCTGTGGGACGAGAAAGCAAGAACTGGGACTAGG	781		
QY 1025	ccagaagttcccaactgtgagggggaagagctgggagcaagctcctccctggatccctgt	1084		
Db 782	CCAGAAAGTTCCCACTGTGAGGGGGGAAGAGCTGGGGAAGCTCCTCCCTGGATCCCTAGT	841		
QY 1085	ggaatttgaagaagactatttcttcttctgtgacacaaatgtttaaattgaatttaa	1144		
Db 842	GGATTTTGAAGAAAGATACATATTTTATTTATTTATGTGACAAAATGTTAAATGATATTAAA	901		
QY 1145	gagaaataaactgattctctctc	1168		

	DB	300	TGGCTGGAGAGACCAAAATACAAGCTCCAGCCTTGGGCTACGACCGCCAGATTGG	359
	OY	382	ggaattacagcaacaaaggctggcgtctaactactgtacatgtacagtgcacatttgata	441
	DB	360	GGAATTACAGTCATCAAGGCGTGCGGTCTACTACTGTACTGTACAGTGACATTGTGATA	419
	OY	442	gggaaagcgctgtcacctgaagctgacctgtgtgtgaacgggtgtgtgccttcgcgtg	501
	DB	420	GGGAAAGCGCTGTCTACCTGAAAGCTGGAAGTTGCTGTGTGAACGGTGTGTCGGCCTGCCTG	479
	OY	502	ccttgaagaattcttaagccacaagcagaagctctctctgtggccccagctccgtttgtgcca	561
	DB	480	CCTGGAAGAATTCTCAGCCACGACGACAAAGCTCTCTGGGCCCAAGCTCTTTGTGTCCA	539
	OY	562	gggtctggcgctgtttgcgcgtcggcgcgaagggtctctcccttcgtagtcgcacaccctcgtg	621
	DB	540	GGTGTCTGGGCGCTGTTGGCCGCTCGGGCCA-GGCTTTCCTTGGATCCGACCCTCCCTG	598
	OY	622	ggcatcatttaagctgtccccttctaactacttgyagactctttaagttcaatcagaag	681
	DB	599	GGCATCATCTTAAGCGCTG-CCCCCTCTTAACCTAACCTTGGACTCTTCAAGTTCACTGAGG	657
	OY	682	ggcct c 686 	
	DB	658	GGCCT 662	
RESULT	3			
LOCUS	B1871711	731 bp	mRNA	linear EST 11-OCT-2001
DEFINITION	603395825F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405478 5', mRNA sequence.			
ACCESSION	B1871711			
VERSION	B1871711.1	GI:16045386		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rgs@bbs-femail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLM12034 row: b column: 07 High quality sequence stop: 728. Location/Qualifiers 1..731 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5405478" /clone_lib="NIH_MGC_90" /lab_type="adenocarcinoma, cell line" /lib_host="DH10B (phage-resistant)" /note="Organ: liver; Vector: PCMV-SpOx6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."			
FEATURES	source			
BASE COUNT	125 a	242 c	228 g	136 t
ORIGIN				

	DB	659	TCACGAGGGGCCCTCGTCTCCCCG	683
	RESULT	5		
	LOCUS		AU763237	
	DEFINITION		unr70d09.y1 NCI CGAP Mam3 Mus musculus CDNA clone IMAGE:3155633 5'	
			similar to TR:O54907 O54907 TNF-RELATED WEAK INDUCER OF APOPTOSIS	
			; mRNA sequence.	
	ACCESSION		AU763237	
	VERSION		AU763237.1 GI:7695174	
	KEYWORDS		EST.	
	SOURCE		house mouse.	
	ORGANISM		Mus musculus	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
	REFERENCE		1 (bases 1 to 561)	
	AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
	TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
	JOURNAL		Tumor Gene Index	
	COMMENT		Unpublished (1997)	
			Other_ESTS: unr70d09.xl	
			Contact: Robert Strausberg, Ph.D.	
			Email: cgaaps-remail.nih.gov	
			Tissue Procurement: lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.	
			cDNA Library Preparation: Life Technologies, Inc.	
			cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
			DNA sequencing by: Washington University Genome Sequencing Center	
			clone distribution: NCI-CGAP clone distribution information can be	
			found through the I.M.A.G.E. Consortium/LNL at:	
			image.lnl.gov/image/html/tiresources.shtml	
	FEATURES			
	Source		MGI:1058389	
			Seq primer: -40RP from Gldco	
			High quality sequence stop: 433.	
			Location/Qualifiers	
			1..561	
			/organism="Mus musculus"	
			/strain="129 - C57/B6 - EVBN"	
			/db_xref="taxon:10090"	
			/clone="IMAGE:3155633"	
			/clone_lib="NCI-CGAP_Mam3"	
			/tissue_type="tumor, gross tissue"	
			/dev_stage="10 months"	
			/lab_host="DH10B"	
			/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;	
			Site_2: NotI; Cloned unidirectionally. Primer: oligo dr.	
			library constructed by Life Technologies. Investigators	
			providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH	
			Reference for transgenic model: Xu et al., Nature Genetics	
			22, 37-43 (1999)."	
	BASE COUNT		108 a 158 c 194 g 100 t 1 others	
	ORIGIN			
	Query Match		43.2%	Score 504; DB %; Length 561;
	Best Local Similarity		99.8%; Pred.No. 2.6e-112;	
	Matches 504; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
OY	1	ggtctgagcctggcgccgtgcgtgcgtgcgtgcgttcctctgtctgctggtcaacct	60	
Db				
	57	gctccttgaggccctgggccctggccctggccctggccctgcctgcctgcgtgcgtacgct	116	
OY	61	ggggagcctggggcaaagctgcttcccacagaagccttcacgagagagcctcacagacaaga	120	
Db				
	117	gggagagctggggcaaacgctgctctccccacagagactttcttagagagagctgacacagaga	176	
OY	121	ccgaccggagagcccccctgtaactgtaatcccccacagagagaaagccagagtgtgtacctt	180	
Db				
	177	ccggccggagagcccccctgtaactgtaatcccccacagagaaaagccagagtgtgtacctt	236	
OY	181	cttggaacaactagtccggcctcgagaagaagtgtctctaagaagccggaagcggcgacctg	240	

Db	237	CTGGGACACACTAGTCGCGCCCTCGAAGAAAGTCTCTCTTAAGGCGCGAAGGCGGCGCTCG	296
OY	241	ccgagctatgtcagcccatltagagttcatctctgcgcagagacagatgtagacaagc	300
Db	297	CCGACCTATTGTCAGCCCATTTATGAGGTTCTATCTCTGCGCAGAGACAGGATGAGACACAAGC	356
OY	301	aggctgagatggagaaagfagatgctgtggagaagagccaaatcaacagctcagccctc	360
Db	357	AGGTTGATGGAGACAGTGAAGTGGCTGGGAAGAGACCAAAATCAACACACTCCAGCCCTCT	416
OY	361	ggcgcacagaccgcagatgggaaattacagtlcagagcgtggcctctactacctg	420
Db	417	GGGCTACGACGCCCGCAGATTGGGGATTTACAGTATCAAGGCGTGGGCTCTACTACTCTTA	476
OY	421	ctgtcagtgacactltagatgagaagagcgtgtctacttgaaagctgagctgtgtgtaa	480
Db	477	CTGTGAGGTGCGCACTTTGATGAGGGAAAGCCTGTACTGAAAGCTGGAGCTTGCTGTGAA	536
OY	481	cggctgctgcgcctcgagctgctg	505
Db	537	CGGTGTCTGCGCCCTGCGCTGCTG	561
RESULT	6		
LOCUS	BE628951	533 bp	mRNA linear EST 25-AUG-2000
DEFINITION	un30c03.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone IMAGE:3373444 5' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.		
ACCESSION	BE628951		
VERSION	BE628951.1	GI:9111639	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Enkaiyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 533)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:1083048		
FEATURES	Seq primer: -40RP from Glibco High quality sequence stop: 437.		
SOURCE	Location/Qualifiers 1..533 /organism="Mus musculus" /db_xref="taxon:10090" /clone="IMAGE:3373444" /clone_lib="Soares_mammary_gland_NMLMG" /sex="female (lactating)" /tissue_type="mammary gland" /lab_host="DH10B" /note="Vector: pRTT30-Pac (Pharmacia) with a modified polylinker. 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	91 a 206 c 105 g 131 t		
ORIGIN			
Query Match	42.6% Score 497; DB 10; Length 533;		
Best Local Similarity	99.4% Pred. No. 1.3e-110;		
Matches 530; Conservative	0; Mismatches 0; Indels 3; Gaps 3;		

FEATURES	Source
LOCUS	BG404836
DEFINITION	6024220160F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4527038 5',
ACCESSION	BG404836
VERSION	BG404836.1
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 650)
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@rsf-mail.nih.gov Tissue Procurement: The Cepko Laboratory cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLNL10435 row: h column: 15 High quality sequence stop: 468. Location/Qualifiers 1..650

Site_2: Salt: Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH_MGC Library."

Query Match	41.9%;	Score 489.6;	DB 10;	Length 650;
Best Local Similarity	92.0%;	Pred. No. 8.9e-109;		
Matches 550;	Conservative 0;	Mismatches 44;	Indels 4;	Gaps 3;

QY	291	gagcgaagaagctgtctgagctggagagctgagctgagctggggaagaagccaaatcaacagct	350
Db	25	GACCCATTTCAGAGTGTGGATGGAGAGTAAGTGGCTGGGGAAGACCAAAATCAACAGCT	84
QY	351	ccagccctctgctgcacagaccagatctgggaaatttcaagltcaatcaaggtctgagctct	410
Db	85	CCAGCCCTCTGGCGTACAGACACCCCGAGATTGGGGAATTTCAGTCAATCAGGCTGGGCTCT	144
QY	411	actaacctgactgtcaagltgcacttgaatgaaggaaagctgtcttaacctgaagctgagct	470
Db	145	ACTACTGTACTGTGACGGTGCACCTTTGATGAGGGAAGAAAGCGCTGTCTACCTGAAGCTGAGCT	204
QY	471	tgtctgtgaagctgtgtctgagccctgagctgtccttggaagaattcttaagccacagagca	530
Db	205	TGCTGGTGAACGGTGTGTGGCCCTGGCCCTGGCAAAATTCCTACGACAGCAGCA	264
QY	531	gtctcccttgagcccaagctcgttctgtcaagtgctgtcgtgagctgtctgcctctgagccag	590
Db	265	GCTTCTCTGGGCCCCAGCTCCGTTTGTCGACAGTGTCTGGCTGTGGCGCTGGGCCA	323
QY	591	ggtctccctctgcgatcgcgacacccctcccttggagctcatcttaagctgcacccctctaa	650
Db	324	GGTCTTCCCTTGGATCCGACCCCTCCCTGGGCTCATCTTAAAGCTGCCCTTCTCTAA	383
QY	651	cctactcttgagctcttcaagttcaagtcaatgagggagctgtctctccagatctcttaact	710
Db	384	CTTACTTTGGACCTCTTCAAGTTCACTGAGGGGCGCTTGTCTCCGAGTTCTTAAACTT	443
QY	711	tcccttgagctcagagagcatcaccaacccctccctaccccaaccccaactctccaccccctc	770
Db	444	TCCCTTGCTCCAGAGCATCACCACTGACTACCCACACAGTAACCTCCAGCCACATC	503
QY	771	gctgtcctcttggltccagltcgtl-ctctcctcaaaagcagcagagctgtgtcaatgt	829
Db	504	GCTGATCCTTGGTCCGAGTCTGCTGCTCATCAAAAGCAGACAGACTTGTACACATGTA	563
QY	830	tccatccacagacgtatcctctgtctcttcaatccatccatccacacaactatcca	887
Db	564	ATCAT--CCAAAGAGATTAAGTACTTACTTAAATGACCATGCAAAACAATATACACA	619
RESULT 8			
AW917574			
LOCUS	AW917574	564 bp	mRNA linear EST 25-MAY-2000
DEFINITION	EST1348878 Rat gene index, normalized rat, norvegicus, Bento Soares		
ACCESSION	R130 norvegicus cDNA clone R01EF49 5' end, mRNA sequence.		
VERSION	AW917574		
KEYWORDS	AW917574.1 GI:8083328		
SOURCE	EST.		
ORGANISM	Norway rat.		
REFERENCE	Rattus norvegicus		
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;		
	Rattus.		
AUTHORS	1 (bases 1 to 564)		
	Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J.,		
TITLE	Kerlavage,A.R. and Adams,M.D.		
	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
JOURNAL	Gene Index		
	Unpublished (1998)		
COMMENT	Contact: Lee, NH		

Db	374	CTGGGAGATGGTGTCTTGCCCTCGGCTGCCTGGAGGAATTCTCAGCCACTTCGGCGAGT	433
Oy	533	tctcctgagcccaagccttcggttgcacagtgtctggcgtgttcgcgctcgccagg	592
Db	434	TCCCTCGGGGCCACACCTCTGGCCAGGAGTGTGGGCTGTGGCCCTGGGGCCAGG	493
Oy	593	tcttccttgatccgaccacctccccgggcttaactaaagctgccccttcctaac	652
Db	494	TCTCTCCCTCGGGATCCGACACCTCCCTCTGGGCCCATCTCAAGGCTGCCCTTCTCAC	553
Oy	653	tacttggacactcttcaagttoactgtaggggacctgctctccccaattccttaacttc	712
Db	554	TACTTCGGAGCTTCTCAGATTACAGAGGGGCCCTGGTCTCCTCCGCACTGCCAAGCT	613
Oy	713	cctgtgccaggagcatcaaccacactccctaccccccaactctccaccctc-g	771
Db	614	GCCGCGCTCC-----CCTGCAGACGCTCTGGGACACCCTGCTCCCTGCCCCACCTCAG	668
Db	772	ctgcctccttgtcacagtcctgtctctcc--tcaaaggcagccagagctgttacaatgt	829
Oy	669	CGGCTTTTGCTCCAGACCTGCCCCCTCTCTAGAAGCTGCTGGGCCCTGTTCACAGTGT	728
Oy	830	tccattcc	837
Db	729	TTCCATCC	736
RESULT	10		
LOCUS	B1766766	834 bp	mRNA linear EST 25-SEP-2001
DEFINITION	6030568661 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5'		
ACCESSION	B1766766		
VERSION	B1766766.1	GI:15758344	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-f@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM1517 row: c column: 18 High quality sequence stop: 772. Location/Qualifiers 1..834 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="5206217" /clone_lib="NIH_MGC_122" /lab_host="DH10B" /note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung; 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note this is a NIH_MGC Library."		
FEATURES	Source		
RACE COUNT	154 a	287 c	223 g 170 t

ORIGIN	Query Match	40.5%	Score 473;	DB 10;	Length 834;
	Best Local Similarity	78.0%	Pred. No. 1,1e-104;		
	Matches 653; Conservative	0;	Mismatches 160;	Indels 24;	Gaps 6;
Qy	66	ctcaagagaagctctcaacagagagacgcgcggagagcccccgaactgaatcccccagacag	155		
Db	1	CCCCAGAGAGAGCTGGTGGCAGAGAGACACAGACCCTCGGAACTGAATCCCGACAG	60		
Qy	156	aggaagacagagatggtgtacattctcttgaaacaactagtcgcgcctcgtgaagaagtgtc	215		
Db	61	AAGAAAGCCAGAGATTCCTCGCCCTTCCGTAACCGACTGATTCGGCTGCGAGAACTGCAC	120		
Qy	216	ctaaagccggaagagcgcgccctcgccagagctatgtcacaccatatagtaggtltaactc	275		
Db	121	CTAAAGGCCGGAAAAACAGGGGCTCGAAGAGCGATGCGACCCATATTAAAGTTCATCCAC	180		
Qy	276	ggcc aggaacatgagatggagagcaacagcaagtggtggaagagtgagcgtcgggaagag	334		
Db	181	GACCTGTGACAGAGAGAGGAGCGAGGACAGTGTGGAGCGGAGACAGTGAAGTGGCGGGAGGAA	240		
Qy	335	accnaaatcaacagctccagccctctggtcgtacagacgcagatgtgggaatttaacgtc	394		
Db	241	GCCAGAAATCAACAGCTCCAGCCCTTGCGGCTTAACAACCCCGAGATGGGGAGATTTAAGTC	300		
Qy	395	atcaaggtctgggtctactactctgtactgtcagtgacacttgatgaaaggaagctgtc	454		
Db	301	ACCCGGGCTGGGCTCTACTACTGTAAGTCAAGTGACGATCTTGATGAGAGGGAAGGCTGTC	360		
Qy	455	tactgtgaagctgtgaactgtctgtgaaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	514		
Db	361	TACCTGAAGCTGTGACTTCTGTGTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	420		
Qy	515	tcaagcacaagagcaagagcttcccttggtggcccaagctcgtttgtgcagaagtgtcgtgctg	574		
Db	421	TCAAGCACTGGGCGCCAGTTCCCTCGGGGCCCAAGTCCGCTCTGCGAGAGTGTGTGGCTG	480		
Qy	575	ttgcgcgtcgtgcgcgcagaggtctctccctctggaatcgacacccctcccttggtctatcttaag	634		
Db	481	TTGGGCGCTGGGCGCA -GGTCTCTCCCTGGGAGATCGCACCCCTCCCTGGGCCCATCTCAAG	539		
Qy	635	gctgcgcctctcttaactactcttggaactcttcaagtltaactgaagggcctgtctctcc	694		
Db	540	GCTGCGCCCTTCTCTACCTACTTTCGAGACTCTTCCAGTGTCACTGAGGGGCGCTCGTCTTC -	598		
Qy	695	cagatctcttaactcttccctgtgtgcgcagagagcatcaacaacacccctcct - - - acccaac	750		
Db	599	-----CCACACGTGTCGCCAGGCTGGCGGCTCCCTCGACAGCTCTCTGGGACACCCGGT	652		
Qy	751	cccacatccatcaacccccctgcgtctctctgttgcagtlcgtctctctc - - - tcaagagca	808		
Db	653	CCCTCTTGTGCCCAACCTTCAGCGGCTCTTGTGTCCAGAACGCGCCCTCCCTTAAGGCT	712		
Qy	809	gccagagctgtgtacatgtltaactc - - - - - acagaagtltaactcgtctctctc	859		
Db	713	GCTTGGGCTGTTCACGTGTGTTCATCCACATTAATTAAGTATTCCACACTTATCTT	772		
Qy	860	taacttccatctccacacaaataatcaacacttaactagcttccaaagccctactctat	916		
Db	773	AAAAATTTCCCCCAAGGCCCACTTCACATCTATTAGTCCCAATTCCTCGACCTTT	829		
RESULT 11					
LOCUS	B1596681	828 bp	mRNA	linear	EST 07-SEP-2001
DEFINITION	60324254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',				
ACCESSION	B1596681				
VERSION	B1596681.1	GI:15489620			
KEYWORDS	EST.				
SOURCE	human				

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 828)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1M1722 row: k column: 13
High quality sequence stop: 776.
Location/Qualifiers

FEATURES

source
1. 828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/clone_id="NIH-MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 155 a 278 c 223 g 172 t
ORIGIN

Query Match 39.1%; Score 456.4; DB 10; Length 828;
Best Local Similarity 82.6%; Pred. No. 1.2e-100;
Matches 550; Conservative 0; Mismatches 106; Indels 10; Gaps 2;
QY 137 gaactgaatcccgagagagagagagatgtgtactctcttggaacaactgtc 196
DB 44 GAATCTGAATCCCGACACGAGAAAGCCAGATCTCGCTTCGTAACCGACTAGTT 103
QY 197 cggcctcgaaagtgctcctaagcgcgagcgccctcgcgagctatcgagcc 256
DB 104 CGGCTTCGCAAGTGCACCTAAAGGCCGGAACACGCGGCTCGAAGAGCGATCGACCC 163
QY 257 cattaatgagttcactcctcgcgagagagatggaagcaagaagtgtgagtggaca 316
DB 164 CATATATGAAGTTTCATCCAGCAGCTGACAGACGAGCGACGACGAGGTGTGACGGGACA 223
QY 317 gtgagtgctgctggaagagagcaaaatcaaacgctcagccctctgagctacagccagc 376
DB 224 TTGAGTGCTGCGAGAGCCAGATACACAGCTCCACCTCTGCGCTACACCGCCAG 283
QY 377 attgggaatttaacagatcaaggctggtgctcactactcgtactgtcagagtgcaattt 436
DB 284 ATCGGGAGTATTATAGTACCCGGGCTGGCTTACTACTGCTACTGACAGTTCATT 343
QY 437 gatgagggaaaagctgtctactctgaagctgtgactgtgtgaaagtggtgtgacctg 496
DB 344 GATAGGGGGAAGGCTGCTACTACTGAAGCTGACCTGCGTGGTGTGATGTGCTGCGCCTG 403
QY 497 cggcgcctggaagatctcagccagcagcagcagcagctctctggtggcccaactcggtttg 556
DB 404 CGCTGCTCGGAGGAATTCACGCCACCTGCGGCGCCCGCCACCTCGGCTTC 463
QY 557 tgcgaagtgtctggtgtgtgctgctgctggtgaggggtctctccttcgagatcgcaacctc 616

DB 464 TGCCAGGTCGTGGGCTGTGTGGCCCTGCGGCGAGGGTCTTCCTCGGAGTCGACCCCTC 523
QY 617 ccttgagctacttaagctgtccccctctactaactacttgaactcttcaagttcac 676
DB 524 CCTGGGGCCATCTCAAGGCTGCCCCCTTCCTCACTTCTCGGACCTTCCTCAGGTTTCC 583
QY 677 tgaaggagctgtgctctccagagatctcctaacttcccttgctcgaagagatcacaca 736
DB 584 TGAGGGGCTCTGCTGTCTCCCG-----CACTCTCCAGGCTCGGCTCCCTCGACCA 636
QY 737 cctccct---accacaccactctccaccctcctgctgctctgtgtccagtcctgt 793
DB 637 GCTCTGTGGGACCCGGTCCCTCTGTGGCCACCCCTCAGCAGCTCTTCTCGACAGCTTC 696
QY 794 ctctcc 799
DB 697 CCTTCC 702

RESULT 12
AA221610 471 bp mRNA linear EST 13-FEB-1997
LOCUS
DEFINITION my18d09.t1 Barstead mouse heart MPLRB3 Mus musculus cDNA clone
IMAGE:696209 5', mRNA sequence.
ACCESSION AA221610
VERSION AA221610.1 GI:1840863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 471)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Stepien,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:429769
Seq primer: -28m13 rev2 EF from Amersham
High quality sequence stop: 301.
Location/Qualifiers

FEATURES

source
1. 471
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:696209"
/clone_id="Barstead mouse heart MPLRB3"
/sex="mixed"
/tissue_type="heart"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Organ: heart; Vector: pT773D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTATCAATCTCAAGGAGGAGCGGCGCCCTTTTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CTTGAATTCGGTAC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 91 a 141 c 128 g 111 t
ORIGIN


```

Db      285  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CACTAGCTCCCAAAGCCCCCTACTTATCCCTGACTCCCCACCACCTCACCCGACCCTG 226
Oy      951  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ttlattgacttcttgaccaggcaactgagatggctgacctggtgaggaagccagag 1010
Db      225  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTTATTGACTTTGTGCACACGACCTGAGATGGCTGGAACCTGTGCGCAGGAGCCAGAG 166
Oy      1011 aacctggagactagccagaagttcccaactglgaaggaggaaagctggggacaagctcct 1070
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      165  AACCTGGGACTAGGCGCAGAGTTCCTCACTGTGAGGGGGAAGAGCTGGGGACAACCTCCT 106
Oy      1071 cccctggatcccttgagatlttgaaagatactatlttatactatctgacacaaatgct 1130
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      105  CCTGTGATCCTCTGTGATTTGAAAAGATACATATTTTATTATTTATGTGACAAAAATGTT 46
Oy      1131 aaatgatattcaagagaataaatacatga 1159
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45  AAATGGATTTTAAAGAGAAATTAATCATGA 17

```

Search completed: June 22, 2002, 10:48:59
 Job time: 18054 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 10:01:03 ; Search time 328.09 Seconds

(without alignments)
6112.213 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168
Sequence: 1 ggtgctgagccctggcctg.....ataatcatgattctcttc 1168

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

tal number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.GeneSeq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1168	100.0	1168	19	AAV18599
2	699.4	59.9	701	20	AAAX23425
3	628.6	53.8	1353	21	AAA49717
4	628.6	53.8	1421	20	AAAX56000
5	614.6	52.6	1373	19	AAV18600
6	597.8	51.2	1236	19	AAV47613
7	597.8	51.2	1236	22	AAAD04350
8	522.8	44.8	1030	20	AAAX23424
9	498.8	42.7	898	22	AAAS03964

10	82.8	7.1	282	16	AAI22190	Human gene signatu
11	46.6	4.0	53526	19	AAI94101	Human PKD1 gene.
12	46.6	4.0	53577	17	AAI18551	Human polycystic k
13	46.6	4.0	53577	19	AAI94108	Human PKD1 locus b
14	45	3.9	1337	20	AAI217263	Human gene express
15	44.8	3.8	1000	21	AAA02484	Human colon cancer
16	43.6	3.7	16167	24	ABJ33083	Human immune syste
17	43.6	3.7	16167	24	ABJ34529	Human metastasis a
18	43.2	3.7	720	22	AAI05001	Human CDNA clone (
19	43.2	3.7	2260	22	AAH18456	Human CDNA sequenc
20	43.2	3.7	2272	22	AAH34689	Human DNA for a no
21	43.2	3.7	2272	22	AAH34690	Human DNA for a no
22	42.8	3.7	1166	20	AAI19440	M. tuberculosis an
23	42.8	3.7	1166	20	AAI19228	M. tuberculosis re
24	42.6	3.6	320	21	AAI38183	Primer used in the
25	42.4	3.6	1286	24	ABJ99656	Mouse ischemic co
26	42.4	3.6	10732	21	AAI10594	Gene encoding a su
27	41.8	3.6	14006	24	ABJ33959	Human immune syste
28	41.4	3.5	320	21	AAI38186	Primer used in the
29	41.4	3.5	1235	20	AAI16147	Human gene express
30	41	3.5	1593	21	AAI02504	Human colon cancer
31	40.8	3.5	1336	20	AAI26779	Human gene express
32	40.8	3.5	1336	21	AAI01591	Human colon cancer
33	40.6	3.5	4108	18	AAI70123	DNA encoding beta-
34	40.4	3.5	50	20	AAI56002	Human tumour necro
35	40.4	3.5	50	21	AAI49732	Human PRO207 DNA p
36	40.4	3.5	318	21	AAI38184	Primer used in the
37	40.4	3.5	320	21	AAI38185	Primer used in the
38	40.4	3.5	6334	24	ABJ33212	Human immune syste
39	40	3.4	1036	20	AAI19421	M. tuberculosis an
40	40	3.4	1036	20	AAI19209	M. tuberculosis re
41	39.2	3.4	1213	24	ABJ01538	Murine apoptosis r
42	39.2	3.4	1517	21	AAI72028	CDNA encoding huma
43	39.2	3.4	2188	20	AAI27506	Human ovarian tumo
44	39	3.3	410	22	AAI182055	Human polynucleoti
45	38.8	3.3	766	22	AAI192737	Human polynucleoti

ALIGNMENTS

RESULT 1	
AAV18599	AAV18599 standard; CDNA; 1168 BP.
XX	AAV18599;
AC	21-JUL-1998 (first entry)
XX	
DT	Mus musculus tumour necrosis factor related ligand (TRELL) gene.
XX	
DE	TRELL: tumour necrosis factor related ligand; tnfr; treatment;
XX	KW cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection; ds.
KW	
XX	
OS	Mus musculus.
XX	
FH	Location/Qualifiers
FT	2..679
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
XX	
PN	W09805783-A1.
XX	
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
PR	18-OCT-1996; 96US-0028515.
XX	
PA	(BIOT) BIOGEN INC.

PA (UIGE-) UNIV GENEVA FACULTY MEDICINE.
 XX Browning JL, Chicheportriche Y;
 XX WPI: 1998-145619/13.
 DR P-PSDB: AAW47524.
 XX
 PT Tumour necrosis factor related ligand - useful for, e.g. treating
 XX cancer, auto-immune disease and immune responses to tissue grafts
 PS
 XX Claim 2: Pages 45-46; 69pp; English.
 CC The sequence is that encoding mouse tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. The DNA sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 SQ Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other:
 Query Match 100.0%; Score 1168; DB 19; Length 1168;
 Best Local Similarity 100.0%; Pred. No. 6,3e-307;
 Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 cggctgctgcgcctgcgctcgtcgtggaagaaattcacaagacaagcagaagctctcctgg 540
 Qy 541 gccccagctccgcttctgtccaaagtgtctgggctgttcgcgcgtcgcgcagaggtctccct 600
 Db 541 gccccagctccgcttctgtccaaagtgtctgggctgttcgcgcgtcgcgcagaggtctccct 600
 Qy 601 tcggaatccgacccctcccttggtgcataccttaaggctgccccctcctaactcttgg 660
 Db 601 tcggaatccgacccctcccttggtgcataccttaaggctgccccctcctaactcttgg 660
 Qy 661 actcttcaagttaagtgaggcgcttgcctcccaagattcccttaacttccctggctc 720
 Db 661 actcttcaagttaagtgaggcgcttgcctcccaagattcccttaacttccctggctc 720
 Qy 721 caggagcatcaccaacccctccctaccccaaccctccctccacccctgcgtcctt 780
 Db 721 caggagcatcaccaacccctccctaccccaaccctccctccacccctgcgtcctt 780
 Qy 781 ggtcagctcgtctctccccaaggcagcagagctgttcacatgtttccatccaca 840
 Db 781 ggtcagctcgtctctccccaaggcagcagagctgttcacatgtttccatccaca 840
 Qy 841 gacgtatcctgtctctcttcaacatccatccaccacaacatacctaactaactagctcc 900
 Db 841 gacgtatcctgtctctcttcaacatccatccaccacaacatacctaactaactagctcc 900
 Qy 901 caaagccctacttaccctcgtactccccaaccctccccaaccagctgttatagact 960
 Db 901 caaagccctacttaccctcgtactccccaaccctccccaaccagctgttatagact 960
 Qy 961 ttgtcaccaagcactagatgagctgtgacccctgtgcaagaaagcagaaagcctggagc 1020
 Db 961 ttgtcaccaagcactagatgagctgtgacccctgtgcaagaaagcagaaagcctggagc 1020
 Qy 1021 taggcagaagcttcccaactgtgaggggaaagcttgggaaagctctcccttgatcc 1080
 Db 1021 taggcagaagcttcccaactgtgaggggaaagcttgggaaagctctcccttgatcc 1080
 Qy 1081 ctgtgattttaaagaatactatttatacttatacttgtagaanaatgttaagtatat 1140
 Db 1081 ctgtgattttaaagaatactatttatacttatacttgtagaanaatgttaagtatat 1140
 Qy 1141 taaagaaataataatgatctcttc 1168
 Db 1141 taaagaaataataatgatctcttc 1168

RESULT 2
 AAX23425
 ID AAX23425 standard; DNA; 701 BP.
 XX
 AC AAX23425;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Mouse TNRL3 DNA.
 XX
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; mouse; ss.
 XX
 OS Mus sp.
 XX
 FH key Location/Qualifiers
 FT 1..636
 FT /*tag= a
 FT /product= "TNRL3"
 XX
 PN W09911791-A2.
 XX 11-MAR-1999.
 PD

PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities

PS Example VII: Fig 13A: 156bp: English.

CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.

XX Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match 44.8%; Score 522.8; DB 20; Length 1030;
Best Local Similarity 81.6%; Pred. No. 9e-132;
Matches 643; Conservative 0; Mismatches 137; Indels 8; Gaps 3;

QY 53 gtcaagctggagagctggagcaagctgtctgcccagagcctctcagagagagctgaca 112
DB 1 gtcaagctggagagctggagcaagctgtctgcccagagcctctcagagagagctg 60
QY 113 gcaagagagcgcggcgagccctcgaactgaatccccaagagagagagagagagctgtg 172
DB 61 gcagct 120
QY 173 gtacattctcttgacaactagctccgagagagagagagagagagagagagagagctg 232
DB 121 gcgactctctgagctg 180
QY 233 cggactcgcagagctatgagagagagagagagagagagagagagagagagagagctg 292
DB 181 cggactcgcag 240
QY 293 gcaagct 352
DB 241 gcgagct 300
QY 353 agcactcgcag 412
DB 301 agcactcgcag 360
QY 413 tactctgactgag 472
DB 361 tactctgactgag 420
QY 473 ctgag 532
DB 421 ctgag 480
QY 533 tactctgag 592
DB 481 tactctgag 540
QY 593 tactctgag 652
DB 593 tactctgag 652

DB 541 tctcctctgag 600
QY 653 tactctgag 712
DB 601 tactctgag 660
QY 713 cctgag 771
DB 661 ggcgag 715
QY 772 ctgctctctgag 829
DB 716 cctgctctctgag 775
QY 830 tctcctctgag 837
DB 776 tctcctctgag 783

RESULT 9

AA03964
ID AA03964 standard; DNA; 898 BP.

AA03964;

26-SEP-2001 (first entry)

Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.

XX TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
XX ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
XX retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
XX rheiosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
XX corneal graft neovascularisation; psoriasis; metastatic condition;
XX malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
XX preneoplastic condition; myocardial angiogenesis; wound granulation;
XX scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
XX atherosclerotic plaque neovascularisation; coronary atherosclerosis;
XX peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
XX fusion protein.

XX Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT CDS 52..873

FT /tag= a

FT /product= "Fusion protein comprising a growth hormone

FT leader, a leucine zipper multimerisation

FT domain, and human TWEAK extracellular

FT domain"

WO200145730-A2.

XX 28-JUN-2001.

XX 19-DEC-2000; 2000MO-US34755.

XX 20-DEC-1999; 9905-0172878.

XX 10-MAY-2000; 2000US-0203347.

XX (IMMV) IMMUNEX CORP.

XX WILEY SR.

XX WPI; 2001-417975/44.

XX P-PSDB; AA003499.

XX Modulating angiogenesis in a mammal for treating diseases mediated by
XX angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX peripheral tissue, by administering antagonist or agonist of TWEAK
XX receptor

PS Example 1: Page 39-40; 46pp; English.
 CC The sequence represents a DNA from the expression vector
 CC pDC409-lz-TWEAK, which encodes a fusion protein comprising a growth
 CC hormone leader, a leucine zipper multimerisation domain, and the
 CC extracellular domain of human TWEAK. The fusion protein was used in
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones
 CC from a COS cell human cDNA library. The TWEAK protein is a
 CC member of the tumour necrosis factor (TNF) family and induces
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
 CC used in the treatment and diagnosis of human disease. The disorders
 CC mediated by angiogenesis include ocular disorders characterised by ocular
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
 CC rubellos, uveitis, macular degeneration and corneal graft
 CC neovascularisation, and inflammatory diseases such as arthritis,
 CC rheumatism and psoriasis. Other treatable diseases include malignant and
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
 CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
 CC scleroderma, vascular adhesions, atherosclerotic plaque
 CC neovascularisation, telangiectasia, wound granulation, coronary
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.
 CC
 XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;
 SQ

Query Match 42.7%; Score 498.8; DB 22; Length 898;
 Best Local Similarity 87.0%; Pred. No. 2.8e-125;
 Matches 548; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

OY 56 agcctgggagctggtgcaacgctgtctgcccagagacccctccagagagagctgacagca 115
 DB 250 agtttggtggagcgccgctgctgtcccccagagagcagctgagagctgtgtgca 309
 OY 116 gagagccgagagagcccttgaactgaatcccccagagagagagagcagagatgtgta 175
 DB 310 gagagagagcagagccgctggaactgaatcccccagagagagagagcagagatctctg 369
 OY 176 ccttcttggaacactagctccgctcgaagaatgtccctcaaaagccgagagcgcg 235
 DB 370 ccttccctggaacagctagctcgcctcgcagaaagtcacccaagagccggaacaacag 429
 OY 236 cctcgcagagctatgtcagcccatatagatgtcctcgcgcagagagatgtgagca 295
 DB 430 gctcgaagagcagctcgcagccattatgagttccacgcagcctcgagagagcgagcg 489
 OY 296 cagaagagtggtgagagagagctggtgctggtgagagagacccaataacagctccagc 355
 DB 490 cagcagagtggtgagagagagctggtgctggtgagagagacccaataacagctccagc 549
 OY 356 cctcgcctacagcagcagatltggggaattacagatcaatcagagctgtgcttactac 415
 DB 550 cctcgcgtacacagcagatcgggagattatagtaaccgggctgtgcttactac 609
 OY 416 ctgtactgtcaggtgcaacttgaatgaggaagagctgttactctgaagctgtgactgtcg 475
 DB 610 ctgtactgtcaggtgcaacttgaatgaggaagagctgttactctgaagctgtgactgtcg 669
 OY 476 ggtgaagagtggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 535
 DB 670 gttgagatgtgtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 729
 OY 536 cctggtgcccagcagctcgtgtgtgctgctgctgctgctgctgctgctgctgctgctgct 595
 DB 730 cctggtgcccagcagctcgtgtgtgctgctgctgctgctgctgctgctgctgctgctgct 789
 OY 536 tcccttgcagatcgcagccctccctgtggtcactctaaagctgtgccccttctaaactac 655
 DB 790 tcccttgcagatcgcagccctccctgtggtcactctaaagctgtgccccttctaaactac 849
 OY 656 ttgtgaccttctcaagtctcaactgagggcc 685

DB 850 ttcgagctcttcaggttcaactgagcgcc 879

RESULT 10
 ID AAT22190 standard; cDNA to mRNA; 282 BP.
 XX AAT22190;
 AC AAT22190;
 XX 27-AUG-1996 (first entry)
 DT 27-AUG-1996 (first entry)
 XX Human gene signature HUMGS03761.
 DE
 XX Gene signature: messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09514772-A1.
 XX
 PD 01-JUN-1995.
 XX
 PF 11-NOV-1994; 94WO-JP01916.
 XX
 PR 12-NOV-1993; 93JP-0355504.
 XX
 PA (MATSU/) MATSUBARA K.
 PA (OKUBO/) OKUBO K.
 PI Matsubara K, Okubo K;
 XX
 DR WPI: 1995-206931/27.
 XX
 XX

Identifying gene signatures in 3'-directed human cDNA library - e.g.
 for diagnosis of abnormal cell function, by preparing cDNA that
 reflects relative abundance of corresp. mRNA in specific human
 tissues
 PS Claim 1; Page 1067; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in AAT19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.

Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;

Query Match 7.1%; Score 82.8; DB 16; Length 282;
 Best Local Similarity 80.6%; Pred. No. 1.5e-12;
 Matches 158; Conservative 0; Mismatches 26; Indels 12; Gaps 5;

OY 969 cagcagctgagatgtgctgagctgtg-gtggcagagagcagcagcagcagcagcagcagc 1027
 DB 88 cagcagctgagatgtgctgagctgtg-gtggcagagagcagcagcagcagcagcagcagc 147
 OY 1028 gaagttcccaactgtgaggggagagagcgtgggagagagcagcagcagcagcagcagc 1083
 DB 148 ggaagttcccaactgtgaggggagagagcgtgggagagagcagcagcagcagcagcagc 206
 OY 1084 tggatttgaagaa--agatactatttattatattgtgacaaatgtt---aaatgga 1137

DB 627 NNCNNCCNCNCNCNCNCNCNCNCCTCTCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNC 568
QY 809 gccagagctgttcaatgtttccattccacagacgtatccttgccttcttaacatccc 868
DB 567 CNC 508
QY 869 atcccaaccagaactatccacactcaactagcttcccaaaagccctacttattccctgaactccc 928
DB 507 CGCCNC 448
QY 929 caaccactcaccggag 944
DB 447 CCCCCCNCNCNCNCNC 432

Search completed: June 22, 2002, 11:53:24
Job time: 6741 sec

OM of: US-09-245-198a-2 to: N_Geneseq_032802:* out_format : pfs
Date: Jun 22, 2002 2:32 PM
About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/gen2.1/USPTO.spool/US09245198/runat_21062002_184008_9995/app_query.fasta_1.632
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=trng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPEXT=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YCAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09245198 -ECGN1_1_0 -NCPV=6 -ICPV=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLIPXY -WAIT -THREADS=1

Search information block:

Query: US-09-245-198a-2
Query length: 225
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 287.880000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18599 +		1162.00	1889.93	6.4e-97	11
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23425 +		1089.00	1775.33	1.5e-90	70
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613 +		1020.00	1656.58	6.3e-84	12
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV4350 +		1020.00	1656.58	6.3e-84	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49117 +		1020.00	1655.68	7.1e-84	13
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18600 +		1020.00	1655.54	7.2e-84	14
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV56000 +		1020.00	1655.19	7.6e-84	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV5964 +		951.00	1546.68	8.4e-78	8
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23424 +		951.00	1545.28	1.0e-77	10
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV21473 +		110.50	167.90	0.5271	9
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV41377 +		106.50	156.27	2.34	16
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV41371 +		106.50	156.27	2.34	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV5310 +		106.50	156.27	2.34	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV8714 +		106.50	156.27	2.34	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV5903 +		106.50	156.27	2.34	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV5230 +		106.50	147.65	7.08	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV82935 +		106.50	147.26	7.44	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV64040 +		106.50	147.26	7.44	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV64038 +		106.50	145.34	9.52	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV93866 +		106.50	145.34	9.52	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV64041 +		106.50	141.82	14.96	6
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV93869 +		106.50	141.82	14.96	6
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV82935 +		103.00	159.16	1.62	6
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV69900 +		102.50	155.07	2.73	95
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV39156 +		102.50	155.07	2.73	95
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV29965 +		102.50	155.07	2.73	95
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49024 +		102.50	155.07	2.73	95
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV6986 +		102.50	150.29	5.05	15
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV41489 +		102.50	146.77	7.92	21
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV80224 +		102.50	146.57	8.14	22
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV70284 +		102.50	146.31	8.41	22
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV29966 +		102.50	146.29	8.43	22
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23028 +		99.50	113.45	569.06	38
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV90085 +		99.50	113.45	569.06	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV82934 +		98.00	152.01	4.55	7
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV13915 +		98.00	141.68	15.23	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV106524 +		97.00	133.47	43.64	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAV25519 +		97.00	95.44	5.7e+03	15
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAV84133 +		96.00	132.55	49.13	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62160 +		96.00	119.46	263.11	11

/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62176 - 96.00 96.56 5.0e+03
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH13929 + 95.50 135.65 32.99
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH20349 + 94.00 131.86 53.66
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH98221 + 94.00 129.42 73.33
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH98319 + 94.00 128.85 78.91

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18599

seq_documentation_block:

ID AAV18599 standard; cDNA: 1168 BP.

AAV18599;

21-JUL-1998 (first entry)

Mus musculus tumour necrosis factor related ligand (TRELL) gene.

TRELL; tumour necrosis factor related ligand; tnf; treatment;

cancer; autoimmune disease; immune system; stimulation; suppression;

graft rejection; ds.

Mus musculus.

Key Location/Qualifiers

CDS 2..679

FT /tag= a

FT /note= "tumour necrosis factor related ligand"

PN W09805783-A1.

PD 12-FEB-1998.

PF 07-AUG-1997; 97WO-US13945.

PR 18-MAR-1997; 97US-0040820.

PR 07-AUG-1996; 96US-0023541.

PR 18-OCT-1996; 96US-0028515.

PA (BIOJ) BIOGEN INC.

(UYGE-) UNIV GENEVA FACULTY MEDICINE.

Browning JL, Chicheportiche Y;

WPI: 1998-145619/13.

P-PDB: AAM47524.

Tumour necrosis factor related ligand - useful for, e.g. treating

cancer, auto-immune disease and immune responses to tissue grafts

Claim 2; Pages 45-46; 69pp; English.

The sequence is that encoding mouse tumour necrosis factor related

ligand (TRELL). TRELL or active fragments can be included with a

carrier in pharmaceutical compositions to treat cancer, autoimmune

diseases or immune responses to tissue grafts, or to stimulate or

suppress the immune system. It is useful to screen for TRELL

receptors, by labelling with a detectable label and screening

compositions for binding. Agents interfering with TRELL-receptor

binding can also be screened for, can then be administered,

optionally with interferon- gamma, to induce cell death or

treat, suppress or alter immune responses (especially involving human

adenocarcinoma cells) involving a signal pathway between TRELL and its

receptor. The DNA sequence can be used in gene therapy for

TRELL-related disorders in mammals (especially humans), e.g. tumours,

autoimmune and inflammatory diseases or inherited genetic disorders,

by introducing into cells, and expressing, therapeutically effective

amounts of a vector, e.g. a virus comprising a gene encoding TRELL.

It may also be of use in the preparation of gene probes for

screening natural/synthetic DNAs for TRELL-encoding sequences

and for antisense therapy.

Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;

alignment_scores:

Quality: 1162.00 Length: 225
 Ratio: 5.164 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-2 x AAV18599 ..

Align seg 1/1 to: AAV18599 from: 1 to: 1168

```

1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVa 17
  |||||||
2 GTGCTAGAGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCT 51
  |||||||
17 ValSerLeuGlySerTirPAlaThrLeuSerAlaIngInuProSeGIng 34
  |||||||
52 GGTCAAGCTGGGAGCTGGGCAAGCTGTCTGCCAGGAGCTTCTCAGG 101
  |||||||
34 LuGluLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGln 50
  |||||||
102 AGGAGCTGACAGACAGACCGCGCGAGAGCCCTGAGCTGAACTCCCA 151
  |||||||
51 ThrGluGluSerGlnAspValValProPheLeuGluGlnLeuValArgP 67
  |||||||
152 ACAGAGAAAGCCAGATGTGTGACTTCTTGGAACAACCTAGTCCGGCC 201
  |||||||
67 oArGArSerAlaProLysGlyValArgLysAlaArgProArgArgAlaIle 84
  |||||||
202 TCGAGAAAGTCTCTTAAAGCCGGAAGCGCGCGCTCGCGAGCTATG 251
  |||||||
84 lAlAlaHsTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100
  |||||||
252 CAGCCCATATGATGAGTTCATCCCGGCCAGGACAGATGAGACACAAGCA 301
  |||||||
101 GlyValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerse 117
  |||||||
302 GGGTGGATGGGACAGTGTGCTGGGAAGACCAAAATCAACAGCTTC 351
  |||||||
117 rSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgA 134
  |||||||
352 CAGCCCTCTGCGCTACGACCGCCAGATTTGGGAAATTACAGTATCAGG 401
  |||||||
134 lAgLyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAla 150
  |||||||
402 CTTGGGCTTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
  |||||||
151 ValTyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgGly 167
  |||||||
452 GTCTACCTGAGCTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 501
  |||||||
167 sLeuGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeuAl 184
  |||||||
502 CCGGAAAGAAATTTCTCAGCCACACAGCAAGCTCTCTGGGCCCCACAC 551
  |||||||
184 rGluLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
  |||||||
552 GTTGTGCCAGAGGTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
  |||||||
201 ArgIleArgThrLeuProTirPAlaHisLeuLysAlaAlaProPheLeu 217
  |||||||
602 CGGATCCGAGACCTCCCTGGGCTCATTTAAAGCTGCCCTTCTTAC 651
  |||||||
217 rTyrPheGlyLeuPheGlnValHis 225
  |||||||
652 CTACTTTGAGACTTTCAAGTTTAC 676
  |||||||

```

seq_name: /SID51/gcdata/geneseq/geneseqn-emb1/NA1999.DAT:AA23425

seq_documentation_block:

ID AAX23425 standard; DNA: 701 BP.
 AAX23425:

XX 18-JUN-1999 (first entry)
 DT
 XX
 DE Mouse TNRL3 DNA.
 XX
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; mouse; ss.

OS Mus sp.

XX Key Location/Qualifiers
 FH CDS 1..636
 FT /*tag= a
 FT /product= "TNRL3"

XX WO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNITW) UNITW WASHINGTON.

XX Chaudhary PM;

XX WPI: 1999-205191/17.

XX P-PSDB; AAW93591.

PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities

XX Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

alignment_scores:

Quality: 1089.00 Length: 211
 Ratio: 5.186 Gaps: 0
 Percent Similarity: 99.526 Percent Identity: 99.526

alignment_block:

US-09-245-198a-2 x AAX23425 ..

Align seg 1/1 to: AAX23425 from: 1 to: 701

	15	leuValValSerLeuglySerTPAlaThrLeuSerIaGlnupr	31
	1	CTGGTGTGGTCAGCCTGGGAGCTGGCACCGCTGTCTGCCAAGAGACC	50
	31	osErngIngluLeuThrAlaGluAspArgArgIupProGIuAuA	48
	48	snpProGIthrGluserglAspValValPropheteuGIuInleu	64
	101	ATCCCAACAGACAGAAAGCAGAGTGtGTACTTTCTTGAACAACATA	150
	65	valArgProArgrSerAlaProLysGLArGLysAlaArgProArgr	81
	151	GTCGGGCCTCGAAGAAGTCTCTTAAGCCGGAAGCGCGGCCCTGGCCG	200
	81	galalaialaIaHisTyrgluValHisProArgrProGIuIaAspOlyA	98
	201	AGCTATTGGAGCCCCATTATGAGGTTCATCTCGCCAGCAGAGATGGAG	250
	98	IaGlnIaaglValAspGIyThValSerGIyTrpOIuGIuThrLysIlE	114
	251	CACAAACAGGTGGATGGACAGTAGTGCGGGGAAGAGACCAAATC	300
	115	AasnSerSerProleuarGrTYrrasParGlnIlEglYgluPheThrVa	131
	301	AACAGCTCCAACCCCTTCGGCTACGACCGCCAGATGGGGAAATTTACGT	350
	131	IIlaargalagIeutyrrTyrrLeutyrcysGlnValHisPheaSpOlug	148
	351	CATCAGGCGTGGGCTCTACTACTGTACGTGTCAGGTGCACCTTTGATGAGG	400
	148	IyLySaIaIaIyTrleuLysLLeuAspLeuValasnGIyValIleuaIa	164
	401	GAAGGCTGTCTACCTGAAGCTGGACTGTGGTGAACGGTGTGCTGGCC	450
	165	leuArgcysLeuGIuIupheserAlaThrAlaIaSerSerProGIyPr	181
	451	CTGGCGTGGCTGGAAATTTCTAGCCACAGCAGCAAGCTCTCTGGGCC	500
	181	ogInleuArgLeuCysGlnValserGIyleuleuProleuArgProGIys	198
	501	CCAGCTCCGTTGTGTGCCAGGTGTGGGGCTGTGGCGCTGGCCCAAGGT	550
	198	eSerleuArgIleArgThrleuProTPraIaHisleuysAlaIaPro	214
	551	CTTCCCTTGGATCCGACCCCTCCCTGGGCTCATCTTAAGGCTGGCCCC	600
	215	PheLeuThrTyrrPheGIyleuPheGlnValHis	225
	601	TTCCTAACCTACTTTGGACTCTTCAAGTTCCAC	633
	seq_name:	/SIDSL/gcadata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613	
	seq_documentation_block:	ID AAV47613 standard; cDNA: 1236 BP.	
XX	AC	AAV47613;	
XX	DT	27-Oct-1998 (first entry)	
XX	DE	TNF related endothelium proliferative agent gene.	
KW	ss:	TNF; endothelium proliferative agent; TREPA; wound healing; cancer	
KM	tissue grafting; vascularisation; apoptosis; autoimmune; birth control		
XX	OS	Homo sapiens.	
XX	FT	Key Location/Qualifiers	
FT	CDS	1..750	
TT		/*tag= "TREPA"	
		/product= "TREPA"	

```

XX MO9835061-A2.
XX
XX 13-AUG-1998.
XX
XX PF 12-FEB-1998; 98WO-USO2859.
XX
XX PR 10-FEB-1998; 98US-0021706.
XX
XX PR 12-FEB-1997; 97US-0798692.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX PA
XX
XX PI Wiley SR;
XX
XX DR WPI; 1998-447255/38.
XX
XX DR P-PSDB; AAW29745.
XX
XX PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
XX PT treatment of autoimmune disease, tumours and inflammation
XX
XX PS Claim 11; Page 123-4; 142pp; English.
XX
CC The TNF-related endothelium proliferative agent (TREPA), or its
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
CC promote wound healing or tissue grafting, by promoting vascularisation,
CC also to induce apoptosis for treating cancer and eliminating autoreactive
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
CC TRPA peptides can also be used to target cytotoxic agents or for
CC affinity isolation of the corresponding receptor, the nucleic acid for
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics.
CC Ribozymes, antisense RNA , antibodies or peptides, are used to treat
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
CC vascularisation), inflammation or a wide range of autoimmune conditions,
CC conditions involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and placental
CC formation) or other angiogenic conditions (e.g. ulcers).
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

alignment_scores:
    Quality: 1020.00      Length: 224
    Ratio: 4.744          Gaps: 0
Percent Similarity: 95.982   Percent Identity: 88.839

alignment_block:
US-09-245-198A-2 x AAIV47613 ..

Align seg 1/1 to: AAIV47613 from: 1 to: 1236

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValVa 18
|||||
76 CTGGGCCCTGGGCGCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGC 125
|||
18 lserleuglysertrpalatphrleuseralaglngluproserglnluc 35
|||||
126 CAGTTTGGGAGACCGGCATCGTCGTCCGCCAGACGACTGCCAGAGG 175
|||||
35 lulethrAlaglunspargarlgLuproprogiluleuanproglintr 51
|||||
176 ACCTGGTGCAAGAGAGACGAGCACCGCGTGGAAACGAATCCCCACA 225
|||||
52 glugluserglinaspvalalProphelengluginleuValArpProa 68
|||||
226 GAAGAAGACCAGATCCTGCCCTTTCTTCGAACCGACTAGTTGGCCTCG 275
|||||
68 gArgserAlaProLySglYArGLySalArGrProArGrAlaIleaIaa 85
|||||
276 AAGAAGTGCACTTAAGCCCGAAAACACGGCGCTCGAAGAGCGATGCCAG 325
|||||
85 lahttytGluaValHisProArGrProGlyGlnaspglyalaaglnaagly 101
|||||

```

```
326 CCCATTATGAACTTCATCCAGCAGCTGGACAGGAGGAGCGGACGAGT 375
102 VALASPGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerse 118
376 GTGGAGGAGGAGTGAAGTGGCTGGAGGAAGGAGATCAACAGCTCCAG 425
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
426 CCTCTGCGCTACACACCGCCAGATCGGGGAGTTTATATAGTCACCGGCGT 475
135 lLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
476 GGCTCTACTACTGTACTGTACAGTGCACCTTGTATGAGGAGGAGCTGTC 525
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
526 TACCTGACAGCTGACCTGCTGGTGATGGTGTCTGCGCTGGCGCTGGCT 575
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgL 185
576 GGAGGAATTCCTCAGCCATCGCGGCGAGTTCCCTCGGGCCCGCCAGCTC 625
185 eucysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
626 TCTGCCAGGTGCTGGGCTGTGGCCCTCGCGGAGGAGTCTCCCTGCGG 675
202 lLeuArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTy 218
676 ATCCGACACCTCCCTCCGCGGCGCATCTCAAGGCTGCCCGCTCTCACTA 725
218 rPheGlyLeuPheGlnValHis 225
726 CTTCGGAGCTCTTCACAGTTTCAC 747
seq_name: /SIDSL/gcdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD04350
seq_documentation_block:
ID AAD04350 standard; cDNA; 1236 BP.
AC AAD04350;
XX
XX 04-JUL-2001 (first entry)
XX
XX Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
XX Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
XX TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
XX grafting; vulnerability; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..750
XX FT /*tag= a
XX FT /product= "Human TREPA (TNF related endothelium
XX FT proliferative agent)"
XX
XX US6207642-B1.
XX
XX PD 27-MAR-2001.
XX
XX PF 26-JUN-1998; 98US-0105343.
XX
XX PR 12-FEB-1997; 97US-0798692.
XX PR 10-FEB-1998; 98US-0021706.
XX
XX PA (ABBO ) ABBOTT LAB.
XX
XX PI Wiley SR;
XX
XX WPI: 2001-280760/29.
XX DR P-PSDB; AAB00891.
XX
```

```
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX Example 2; Column 73-74; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,
XX particularly members of tumour necrosis factor (TNF) family molecules
XX designated as TREPA (TNF related endothelium proliferative agent).
XX CC Soluble biologically active TREPA are used to treat TREPA-associated
XX diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX in human for promoting wound healing and for vascularizing grafted
XX CC tissue for successful grafting and to promote tissue grafts.
XX CC The present sequence is a cDNA clone ID #630050 encoding human TREPA.
XX
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:
```

```
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839
alignment_block:
US-09-245-198a-2 x AAD04350 ..
Align seg 1/1 to: AAD04350 from: 1 to: 1236
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValVal 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 125
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
126 CAGTTTGGGAGCGGGGATCGCTGTCGCCAGAGACCTGCCAGAGAG 175
35 lLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGlnThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
176 AGCTGTGGCAGAGAGAGACAGACCGCTCGGAGACTGAATCCCGAGACA 225
52 GluGluSerGlnAspValAlaProPheLeuGluGlnLeuValArgProAr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
226 GAAGAAAGCCAGAGATCTCGCTTCCTGTAACCGACTTAATGGGCTCG 275
68 gArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAla 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
276 AAGAAAGTGCACCTAAAGCGCGGAAACACGCGGCTCGAAGCGCATGGC 325
85 lHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
326 CCCATTATGAACTTCATCCAGCAGCTGGACAGGAGCGAGCGCAGGT 375
102 VALASPGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerse 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
376 GTGGAGGAGGAGTGAAGTGGCTGGAGGAAGCCAGATCAACAGCTCCAG 425
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
426 CCTCTGCGCTACACACCGCCAGATCGGGGAGTTTATATAGTCACCGGCGT 475
135 lLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
476 GGCTCTACTACTGTACTGTACAGTGCACCTTGTATGAGGAGGAGCTGTC 525
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
526 TACCTGAAAGCTGAGCTTGTGGTGATGGTGTCTGCGCTGGCGCTGCT 575
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgL 185
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
576 GGAGGAATTCCTCAGCCATCGCGGCGAGTTCTCTCGGGCCCGCCAGCTC 625
```

185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
626 TCTGCCAGGTGCTGGGCTGTGGCCCTGGCCAGGGGTCTCCCTGGGG 675
202 ILeuGThrLeuProTTPAlaHisLeuLysAlaAlaProPheLeuThrTy 218
|||||
676 ATCCGCACCCCTCCCTGGCCCATCTCAAGGCTGCCCTTCCTCACTTA 725
218 rPheGlyLeuPheGlnValHis 225
|||||
726 CTTCGACTCTTCAGGTTCAAC 747

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA49717

seq_documentation_block:

ID AAA49717 standard; cDNA; 1353 BP.
XX
AC AAA49717;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 cDNA clone DNA30879-1152.
XX
KW PRO207; human; antitumour; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 58..807
FT sig_peptide /tag- a
FT /tag- b
FT mat_peptide 178..804
FT /tag- c
XX
PN WO200037638-A2.
XX
PD 29-JUN-2000.
XX
PE 02-DEC-1999; 99MO-US28565.
XX
PR 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99MO-US05028.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI Napier MA, Pitti RM, Wood WI;
XX
DR WPI: 2000-442668/38.
DR P-SDS: AAY95338.
XX
XX Novel composition to inhibit neoplastic cell growth or for treating
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO219,
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or
PT PRO866
XX
PS Claim 20; Fig 3; 172pp; English.
XX
CC The present sequence is that of cDNA clone DNA30879-1152
CC (AYCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumour necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a

CC foetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

alignment_scores:

Quality:	1020.00	Length:	224
Ratio:	4.744	Gaps:	0
Percent Similarity:	95.982	Percent Identity:	88.839

alignment_block:

US-09-245-198a-2 x AAA49717 ..

Align seg 1/1 to: AAA49717 from: 1 to: 1353

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVal 18
|||||
133 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 182
18 LserLeuGlySerTPAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
183 CAGTTGGGAGCCGCGCATCGCTCCGCCAGAGCGTCCAGAGGAG 232
35 LuLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThr 51
|||||
233 AGCTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 282
52 GluGluSerGlnAspValValProPheLeuGlnGlnLeuValArgPro 68
|||||
283 GAGAGAGCAGAGATCTGCGCTTCCGAAACAGCTAGTGGCGCTG 332
68 ArgSerAlaProLysGlyArgGlyAlaArgProArgAlaAlaAla 85
|||||
333 CAGAGTGACCTAAAGCCGGAACAGCGCTGGAAGAGGATCGCG 382
85 LAlaIstYrGluValHisProArgProGlyGlnAspGlyAlaGlnAcly 101
|||||
383 CCCATTATGAAGTTCATCCACGACCTCGACAGAGGAGCGAGCGAGT 432
102 ValAspGlyThrValSerGlyTyrPheGluThrLysIleAsnSerSerSe 118
|||||
433 GTGGAGGAGGACAGTGTGGCTGGGAGGAGGAGGAGGAGGAGGAGG 482
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
|||||
483 CCTCTGGCGCTACACACCGCAGATGGGGAGTTTATGCAACCGCGCTG 532
135 LylLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
|||||
533 GGCCTACTACTCTGACTGTCAGTGCACATTGTGAGAGGAGGAGGCTGC 582
152 TyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLe 168
|||||
583 TACCTGAAGCTGACTGCTGTGGATGTGTGGCTGGGCTGGGCTGCT 632
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgL 185
|||||
633 GGAGGAATTTTCAGCCACTCGCGGAGTTCCCTCGGCGCCAGCTCCGC 682
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
683 TCTGCCAGGTGCTGGGCTGTGGCCCTGGCCAGGCGCTCCCTCGGGG 732

202	IIleArpghrlreuprOTPrpAlAhlsleuYsAlAAlaProPhelneUrChTy	218
733	ATTCGACACCCCTCCCTGGGCCCACTCAAGGCTGCCCTTCCTCACTCA	782
218	rPhEgLYLeuprheGlnValHis	225
783	CTTCGACTCTTCACAGTTCAAC	804
seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV18600		
seq_documentation_block:		
AC	AAV18600 standard; cDNA; 1373 BP.	
XX		
XX	AAV18600:	
XX		
DE	21-JUL-1998 (first entry)	
XX		
DE	Homo sapiens tumour necrosis factor related ligand (TRELL) gene.	
XX		
XX	TRELL: tumour necrosis factor related ligand; tnf; treatment;	
XX	cancer; autoimmune disease; immune system; stimulation; suppression;	
XX	graft rejection; ds.	
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
XX	CD5	1..852
XX		/*tag= a
XX		/note= "tumour necrosis factor related ligand"
XX	MO9805783-A1.	
XX		
XX	12-FEB-1998.	
XX		
XX	07-AUG-1997;	97MO-US13945.
XX		
XX	18-MAR-1997;	97US-0040820.
XX	07-AUG-1996;	96US-0023541.
XX	18-OCT-1996;	96US-0028515.
XX		
PA	(BIOJ) BIOGEN INC.	
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.	
PI	Browning JL, Chicheportliche Y;	
PI	WPI: 1998-145619/13.	
DR	P-PDB: AAM47525.	
XX		
XX	Tumour necrosis factor related ligand - useful for, e.g. treating	
XX	cancer, auto-immune disease and immune responses to tissue grafts	
XX		
XX	Claim 2; Pages 48-50; 69pp; English.	
CC	The sequence is that encoding human tumour necrosis factor related	
CC	ligand (TRELL). TRELL or active fragments can be included with a	
CC	carrier in pharmaceutical compositions to treat cancer, autoimmune	
CC	diseases or immune responses to tissue grafts, or to stimulate or	
CC	suppress the immune system. It is useful to screen for TRELL	
CC	receptors, by labelling with a detectable label and screening	
CC	compositions for binding. Agents interfering with TRELL-receptor	
CC	binding can also be screened for, can then be administered,	
CC	optionally with interferon- gamma, to induce cell death or	
CC	treat, suppress or alter immune responses (especially involving human	
CC	adenocarcinoma cells) involving a signal pathway between TRELL and its	
CC	receptor. The DNA sequence can be used in gene therapy for	
CC	TRELL-related disorders in mammals (especially humans), e.g. tumours,	
CC	autoimmune and inflammatory diseases or inherited genetic disorders,	
CC	by introducing into cells, and expressing, therapeutically effective	
CC	amounts of a vector, e.g. a virus comprising a gene encoding TRELL.	
CC	It may also be of use in the preparation of prepare probes for	
CC	screening natural/synthetic DNAs for TRELL-encoding sequences	
CC	and for antisense therapy.	

```

SQ Sequence 1373 BP: 247 A: 462 C: 394 G: 270 T: 0 other:
..

Alignment scores:
    Quality: 1020.00      Length: 224
    Ratio: 4.7/4         Gaps: 0
    Percent Similarity: 95.982      Percent Identity: 88.839

Alignment_block:
US-09-245-198A-2 x AAV18600 ..

Align seg 1/1 to: AAV18600 from: 1 to: 1373

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVala 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTG 230
18 IserLeuGlySerThrPalaIhrLeuSerAlaGluProserGlnGlu 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 CAGTTTGGGGAGCGGGCGATCGCTGTCCGCCAGAGAGCTGGCCAGAG 280
35 IueuthrAlaGluAspArgArgGluProGluGluLeuAsnProGlnThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
281 AGCTGGTGGCAGAGAGAGCAGGACCGCGTGGAACTCAATCCACAGAA 330
52 GluGluSerGlnAspValValProPheLeuGluLeuValValArgProAr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
331 GAAGAAAGCCGAGATCCTGGCGCTTCTCTGAACCACTAGTTCGGCGCTCG 380
68 gArGserAlaProLysGlyArgLysAlaArgProArgAlaIleAla 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 CAGAAGCGCACCTAAAGCCGGAAACACGGGCTCGAAGAGGATCGCAG 430
85 IahIstYrGluValHisProArgProGlyGlnAspGlyAlaGlnaGly 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
431 CCATATATGAAGTTCATCCAGACCTGACAGAGACGAGCGCAGGCT 480
102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerSe 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
481 GTGAGCGGCGACAGAGAGTGGCTGGGAGGAAACCAAGATCAACAGCTCCAG 530
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
531 CCGTCGGCGCTACACCGCCAGATCGGGGAGGTTTATGATCACCCGGGCTG 580
135 IYeuuYrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
581 GGGTCTACTACTCTGTACTGTACAGGTGCATTTGATGAGGGGAAAGCTGTG 630
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
168 uGluGluPheSerAlaIhrAlaAlaSerSerProGlyProGlnLeuArgL 185
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
681 GAGGGAATTTCCACCCACTGGCGGAGTTCCTCGGGGCCACAGCTCGGC 730
185 euYsgLInaLserGlyLeuLeuPheProLeuArgProGlySerSerLeuArg 201
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
731 TCTGCCAGGTGTCTGGGCTGTGGCCCTCGCGGCGCAGGGTCTCCCTGGCG 780
202 IleaArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTy 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
781 ATCCGCAACCTCCCTGGGCCATCTCAAGGCGTGGCCCTTCTCTGACSTA 830
218 rPheGlyLeuPheGlnValHis 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
831 CTTCGGACTCTTCAGGTTTCAC 852

seq_name: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:AA556000
seq_documentation_block:
ID AA556000 standard: DNA: 1422 BP.

```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA503964

seq_documentation_block:

ID AA503964 standard; DNA; 898 BP.

xx

```
68 grrgserlarpolysglylrrglvslalaiprccatgacgalaiala 85
|||||
367 CAGAGTCCACCTTAAGACCGGAAACCGGCTCGAAGGCGATCGCG 416
|||||
85 laHsTyGluValHIsProArProClYgInAspGlylaGlnalagly 101
|||||
417 CCCATTATGAAGTTCATCCACGACCTCGACAGGACGAGCCGAGGT 466
|||||
102 ValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerSe 118
|||||
467 GTGGACGGGACAGTGAAGTGGGAGGAAGCCAGATCAACAGCTCCAG 516
|||||
118 rProLeuTrgTyAspArgGlnIleGlyGluPheThrValIleArgAla 135
|||||
517 CCCTTGGGCTACACCCGACGATCGGGAGATTATAGTCACCCGGGCTG 566
|||||
135 lYLeuTyTyTyLeuTyTCysGlnValHIsPheAspGluGlyLysAla 151
|||||
567 GCCTTACTACTGTACTGTCAGGTGCACCTTGTATAGGGGAAGGCTGTC 616
|||||
152 TyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCys 168
|||||
617 TACCTGAAGCTGAGCTGTGCTGGATGGATGCTGCTGACCTGGGCTGCT 666
|||||
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArg 185
|||||
667 GGAGGAATTTCTCAGCCACTGGGGGAGTTCCTCGGGGCCACAGCTCGCG 716
|||||
185 euCysGlnValSerGlyLeuLeuProLeuArgProClYSerSerLeuArg 201
|||||
717 TCTGGCAGAGTGTGGGCTGTGTGGCCCTGGGGCAGGTCCTCCCTGGCG 766
|||||
202 lLeaGTrhrLeuProTrpAlaHIsLeuLysAlaAlaPProPheLeuThr 218
|||||
767 ATCCGACACCTCCCTCGGGCCATCTCAAGGCTGCCCCCTTCTCTACCTA 816
|||||
218 rPheGlyLeuPheGlnValHIs 225
|||||
817 CTTCCGACCTTCCAGGTTCCAC 838
```

seq_name:	/STDS1/gcgdata/geneseq/geneseqn-embL/NA2001A.DAT:AA503964
seq_documentation_block:	ID: AA503964 standard; DNA; 898 BP.
XX	AA503964;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
XX	
XX	TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KW	ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW	retinoblastoma; retinopathy of prematurity; retrorenal fibroplasia;
KW	ruibiosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
KW	corneal graft neovascularisation; psoriasis; metastatic condition;
KW	malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KW	preneoplastic condition; myocardial angiogenesis; wound granulation;
KW	schleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW	atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW	peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
XX	
XX	fusion protein.
OS	Homo sapiens.
OS	Synthetic.
XX	
XX	
Key	Location/Qualifiers
FT	52..873
FT	/tag= a
FT	/product= "Fusion protein compising a growth hormone
FT	leader, a leucine zipper multimerisation
FT	domain, and human TWEAK extracellular
FT	domain"

```

XX  WO200145730-A2.
PN
XX  28-JUN-2001.
PD
XX
XX  19-DEC-2000; 2000WO-US34755.
PE
XX  20-DEC-1999; 99US-0172878.
PR  10-MAY-2000; 2000US-0203347.
XX
XX  (IMMV ) IMMUNEX CORP.
PA
XX
XX  Wiley SR;
PI
XX  WPI; 2001-417975/44.
DR  P-PSDB; AAU03499.
XX
PT  Modulating angiogenesis in a mammal for treating diseases mediated by
XX  angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX  peripheral tissue, by administering antagonist or agonist of TWEAK
XX  receptor
XX

```

Example 1: Page 39-40; 46pp; English.

```

XX  The sequence represents a DNA from the expression vector
XX  PDE409-LZ-TWEAK, which encodes a fusion protein comprising a growth
XX  hormone leader, a leucine zipper multimerisation domain, and the
XX  extracellular domain of human TWEAK. The fusion protein was used in
XX  the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX  from a COS cell human cDNA library. The TWEAK protein is a
XX  member of the tumour necrosis factor (TNF) family and induces
XX  angiogenesis. TWEAKR may therefore be used to screen for and develop
XX  TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
XX  used in the treatment and diagnosis of human disease. The disorders
XX  mediated by angiogenesis include ocular disorders characterised by ocular
XX  neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX  retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX  rubeosis, uveitis, macular degeneration and corneal graft
XX  neovascularisation, and inflammatory diseases such as arthritis,
XX  rheumatism and psoriasis. Other treatable diseases include malignant and
XX  metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX  preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX  scleroderma, vascular adhesions, atherosclerotic plaque
XX  CC  neovascularisation, telangiectasia, wound granulation, coronary
XX  atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
XX  Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:
XX

```

```

XX  Alignment_scores:
XX  Quality: 951.00      Length: 207
XX  Ratio: 4.803        Gaps: 0
XX  Percent Similarity: 95.652  Percent Identity: 88.889
XX

```

alignment_block:

US-09-245-198a-2 x AAS03964 ..

Align seg 1/1 to: AAS03964 from: 1 to: 898

```

19 SerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
   |||||
250 AGTTTGGGGAGCCGCGATCGCTCCGCCAGAGAGCTGCCAGAGGGA 299
   |||||
35 uLeuThrAlaGluAspArgGluProGluLeuAsnProGlnThrG 52
   |||||
300 GCTGGTGGCAGAGAGACGACCGCTGGGAACGAAATCCCAAGACAG 349
   |||||
52 uGluSerGlnAspValValProPheLeuGlnGluLeuValAlaArgProArg 68
   |||||
350 AAGAAAGCCGAGATCTGCGCTTCTCGAAGACGACTAGTTCCGCGCTCGC 399
   |||||
69 ArgSerAlaProGlyGlyArgGlyAlaArgProArgArgAlaAlaAla 85
   |||||

```

```

400 AGAAGTCACACTTAAAGCCGAAAAACAGCGGCTCGAAGAGCATCGCAC 449
85 aHsTYrGluValIHisProArgProGlyGlnAspGlyAlaGlnAlaGly 102
   |||||
450 CCATTTATGAAGTTTCATCCACGACCTGGACAGAGCGAGCGACGAGGTG 499
   |||||
102 aLAspGlyThrValSerGlyTyrPglGluThrLysIleAsnSerSerSer 118
   |||||
500 TGGACGGAGCAGTGAAGTGGCTGGGAGAAAGCCAGAAATCAACAGCTCCAC 549
   |||||
119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
   |||||
550 CCTTGGCGTACACACCGCAGATCGGGAGTTTATGATCCACCGGGCTGG 599
   |||||
135 uLeuTYrTYrLeuTYrCysGlnValIHisPheAspGluGlyLysAlaValT 152
   |||||
600 GCTTACTACTGTACTGTACAGGTGACCTTTGATGAGGGGAAGCGTGTCT 649
   |||||
152 yrlLeuLysLeuAspLeuLeuValAsnGlyValIleuAlaLeuArgCysLeu 168
   |||||
650 ACCTGAAGCTGAGCTGTGGTGAATGTTGCTGCGCCCTGCGCTGCTG 699
   |||||
169 GluGluPheSerAlaThrAlaIAserSerProGlyProGlnIleuArg 185
   |||||
700 GAGGAATTCACAGCCACTGGGCCAGTTCCCTCGGCCCCACGCTCCGCT 749
   |||||
185 uCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 202
   |||||
750 CTCCAGAGTGTCTGGGCTGTGGCCCTCGCGGCAAGGCTCTCCCTCGGA 799
   |||||
202 leArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTYr 218
   |||||
800 TCCGCACTCCCTCCCTGGGCCATCTCAAGGCTGCCCTTCCTCAGACTAC 849
   |||||
219 PheGlyLeuPheGlnValHis 225
   |||||
850 TTCGACTCTTCAGGTTTCAC 870
   |||||

```

seq_name: /SIDS1/gcdata/geneseq/geneseqn-emb1/NA1999.DAT:AA23424

seq_documentation_block:

ID AAX23424 standard; DNA; 1030 BP.

XX AAX23424;

DT 18-JUN-1999 (first entry)

DE Human TNRL3 DNA.

```

XX  Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;
XX  developmental abnormality; gestational abnormality; prostate cancer;
XX  AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX  cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX  apoptosis; human; ss.
XX

```

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..627

FT /*tag= a

FT /product= "TNRL3"

PN WO9911791-A2.

PD 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

PR 05-SEP-1997; 97US-0924634.

PA (UNITW) UNITV WASHINGTON.

XX Chaudhary PM;

XX WPI: 1999-205191/17.
DR P-PSDB: AAM93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT developmental or gestational abnormalities
XX
PS Example VII: Fig 13A; 156pp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SO Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;
alignment_scores:
Quality: 951.00 Length: 208
Ratio: 4.803 Gaps: 0
Percent Similarity: 95.192 Percent Identity: 88.942
alignment_block:
US-09-245-198A-2 x AAX23424 ..
Align seg 1/1 to: AAX23424 from: 1 to: 1030
18 ValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnG1 34
1 GTGAGTTTGGGGAGCCGGGCATCGCTGCCGCCAGAGAGCCTGCCAAGA 50
34 uGluLeuThrAlaGlnAspArgArgGluProGluLeuAsnProGlnT 51
51 GGAGCTGGTGAGAGAGAGACAGACAGCCCTCGAAGTGAATCCCAAGA 100
51 hrtGluLeuSerGlnAspValAlaProPheLeuGlnGluLeuValArgPro 67
101 CAGAGAAAGCAGAGATCGCTTTCGTAACGACACTGTTGGGCT 150
68 ArgArgSerAlaProGlyGlyArgLysAlaArgProArgArgAlaLea1 84
151 CGCAGAGAGTGCACCTAAAGCCGGAACACAGCGCTCGAAGAGCGATCGC 200
84 aAlaHisTyrGlnValAlaHisProArgProGlnAspGlyAlaGlnAlaG 101
201 AGCCATTATATAGATTCATCCACAGCCTGGACAGAGAGCGAGCGAGAG 250
101 LysAlaAspGlyThrValSerGlyTrpGlnGluThrLysLysAsnSerSer 117
251 GTGAGACGGGAGAGTAGTGCTGGGAGGAGAAACCAATCAACAGCTCC 300
118 SerProLeuAlaGlyTrpAspArgGlnLeuGlyLysLysLysValAlaGAl 134
301 AGCCCTCTCGCTACTACCGCCAGATCGGAGAGTTTATAGTACCCCGGCG 350

134 aGlyLeuTyrTrpLeuTyrCysGlnValHisPheAspGlnGlyLysAlaVal 151
1 TGGGCTCTACTACCTGTAAGTGTGACAGTCACTTGTATATAGGGAGAGCTG 400
351 TGGGCTCTACTACCTGTAAGTGTGACAGTCACTTGTATATAGGGAGAGCTG 400
151 aTyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCys 167
401 TCTACCTGAAGCTGACTGCTGGTGATGGTGTGCTGCTGCCCTGGCGTCC 450
168 LeuGlnGluPheSerAlaThrAlaAlaSerSerProGlyTrpGlnLeuArg 184
451 CTGGAGGAATTCACGACATCGGCGCAGTTCCCTCGGCGCCACCTCCG 500
184 gLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuA 201
501 CCTCTGCAAGGTGTCTGGCTGTGGCTGCGGCGCAGAGGTCTCTCTCC 550
201 rGlyLeuArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThr 217
551 GGATCGGACACCTCCCTGGGCGCATCTCAAGGTGCCCTTCTCTCAC 600
218 TyrPheGlyLeuPheGlnValHis 225
601 TACTTCGGAGCTCTTCACAGTTTCAC 624
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL21473
seq_documentation_block:
ID ABL21473 standard; DNA; 978 BP.
XX ABL21473:
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001MO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL1840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 other:

alignment_scores:
 Quality: 110.50 Length: 197
 Ratio: 0.978 Gaps: 10
 Percent Similarity: 57.360 Percent Identity: 24.873

alignment_block:
 US-09-245-198a-2 x ABL21473 ..

Align seg 1/1 to: ABL21473 from: 1 to: 978

```

41 AARGAGLpProgluLeuasnProglInThrgluGluSerGlnAspVa 57
   ::::::::::::::::::::
450 CAAAGAGAGCCCTGCACACCTTCACCA.....CCGTCGCAATGC 490
   ::::::::::::::::::::
57 lvalPro.pheLeuGluInLeuValArgProArgArgSer.....Ala 71
   : ||| ::::::::::::::::::::
491 ATTCCCGCATCGCCACCTCTAGTCGCCAAAGTGATCTCTTTTCA 540
   : ||| ::::::::::::::::::::
72 ProLysGlyArgLysAlaArgProArgArgAlaAlaAlaHisTyrGl 88
   ::::::::::::::::::::
541 GCCAGATCCGAGACTCGAGGCCA.....GCAGCCCATTTCCA 578
   ::::::::::::::::::::
88 uValHisProArgProgluGlnAspGlyAlaGlnAlaValAspGlyT 105
   ::::::::::::::::::::
579 CTTGACACAGCGCGCGTCCACCAAGAAAGTAG...GGCTACCATGAG 625
   ::::::::::::::::::::
105 hValSerGlyTyrGluInThrgLysIleasnSerSerProLeuArg 121
   ::::::::::::::::::::
626 ATATGTACATAGGAATGATMACGAGAAACTCTTATCAGGA...CAC 672
   ::::::::::::::::::::
122 TyrAspArgGlnIleGlyInPheThrValIleArgAlaGlyLeuTyr 138
   ::::::::::::::::::::
673 TTTCAAACCGCGCATGCGCTTGTACGCGTACCAATACAGAGCTATATTA 722
   ::::::::::::::::::::
138 rLeuTyrGlyGlnValHisPheAspGluGlyLysAlaValTyrLeuLys 155
   ::::::::::::::::::::
723 CGATATCGCCAGATVGTCTACCAACATCGCAGAC.....759
   ::::::::::::::::::::
155 euAspLeuValAsnGlyValLeuAla.....164
   ::::::::::::::::::::
760 .....CAGAACGGATTATGCTCTTCAAGAGACACTCCATTC 798
   ::::::::::::::::::::
165 LeuArgCysLeuGluInPheSerAlaThrAlaAlaSerSerProGlyP 181
   ::::::::::::::::::::
799 CTGCAGTCTTGTAAC.....ACGTGCCCAACCAACATCGGCACA 836
   ::::::::::::::::::::
181 oGlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGly 198
   ::::::::::::::::::::
837 TAAAGTGCAACCTGCCACACGAGTGTCTGATCCACCTGGAAAGAAAC 886
   ::::::::::::::::::::
198 eSerLeuArgIleArgThrLeu.....ProThrAlaHisLeuLys 211
   ::::::::::::::::::::
887 AGAGATCCATCTGAAGACATTCACACAGATCGCAATGCAGTTCGCG 936
   ::::::::::::::::::::
212 AlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 244
   ::::::::::::::::::::
937 GAGGGAACACCGAGCTACTTGGCATCTTCAAGGTG 975
   ::::::::::::::::::::
seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV41377
seq_documentation_block:
ID AAV41377 standard; cDNA: 1630 BP.
AC AAV41377;
XX
XX
XX 08-OCT-1998 (first entry)
XX
XX NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.
XX
XX RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
KW

```

KW immune response; inflammatory response; toxic shock; sepsis;
 KW RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.

OS Mus musculus.

XX Key Location/Qualifiers
 FH 3..887
 FT CDS /*tag= a
 FT /product= "murine RANKL (ligand for RANK)"

XX WO9828426-A2.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-US23775.

XX 14-OCT-1997; 97US-0064671.

XX 23-DEC-1996; 96US-0059978.

XX 07-MAR-1997; 97US-0813509.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM, Galibert LJ, Maraskovsky E;

XX WPI: 1998-377657/32.

XX P-PSDB; AAM69956.

PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells

PS Claim 25; Pages 55-57; 80pp; English.

CC This cDNA encodes a murine RANKL, a ligand for the RANK (receptor
 CC activator of necrosis factor-kappa B (NF-kB)) polypeptide. RANK is a
 CC member of the tumour necrosis factor (TNF) family. A soluble RANK
 CC may be used for inhibiting activation of NF-kB, by contacting a cell
 CC expressing membrane-associated RANK with a soluble RANK which binds to
 CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be
 CC used to induce maturation of dendritic cells and enhance their
 CC allo-stimulatory capacity, thereby augmenting an immune response. The
 CC soluble RANK polypeptide composition may also be used for regulating an
 CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists
 CC may be useful in ameliorating negative effects of an inflammatory
 CC response that result from triggering of RANK, e.g. in treating toxic
 CC shock or sepsis, graft-versus-host reactions, or acute inflammatory
 CC reactions. They can also be used in adjunct therapy for disease
 CC characterised by neoplastic cells that express RANK. RANKL polypeptides
 CC can also be used to identify inhibitors of RANK and thus inhibitors of
 CC an inflammatory response, and also for protecting RANK-expressing cells
 CC from the negative effects of chemotherapy or the presence of high levels
 CC of TNF-alpha. The products can also be used for detection and drug
 CC screening.

SQ Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other:

alignment_scores:
 Quality: 106.50 Length: 276
 Ratio: 0.772 Gaps: 13
 Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:
 US-09-245-198a-2 x AAV41377 ..

Align seg 1/1 to: AAV41377 from: 1 to: 1630

```

3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValAlaSe 19
   ::::::::::::::::::::
78 TCCATGTTCTCGCTCGCTCGGCGGAGCTGCGCAAGTGTGCGAG 127
   ::::::::::::::::::::
19 rLeuGlySerThrAlaThrLeuSerAlaGln...GluProSerGlnGluG 35
   ::::::::::::::::::::

```



```

128 CACGCGCTGTCTCTGACTTCTTCCAGGCGCAGATGGATCCTAACAGA.... 173
35 IuleuthralagluAspArg..... 41
174 ....ATATAGAAAGAGCAGCACTCACTGCTTTATAGATTCCTGAGACTC 218
42 ArgGluProProGluLeuAsnProGlnInthrGluGlnSerGlnAspVala 58
219 CATGMAAACGACAGATTGTCAGAGACTGACTCTGGAGATGGAAGACACT 268
58 lPro..... 59
269 ACCTGACTCCTCGAGAGAGATGAACAAGCCTTTCAGGGGCGCTGCACA 318
60 ..PheGluGlnGlnLeuValArgProArg.....SerLapPolyS 73
319 AGCACTGCACACATTGTGGGCGACAGCGCTTCTCAGGAGCTCAGCT 368
74 .....GlyArgLysAlaArgProArgAr 81
369 ATGATGGAAGCGCTCATGCTGTTGATGTGTGCCACGAGCAACCCGTGAGC 418
81 gAlaIleAlaIleHisTyrGluValHisProArgProGlnAspGlyA 98
419 CCAGCCATTGTCACACCTCACCATTCAAGCTGCCAGCATCCCATGGGTT 468
98 IaglnAlaGlyValAspGlyThrValSerGlyTyrGlnGlu..... 111
469 CCGAT.....AAGTCACCTGTCTCTGTGTACCAACATCGAGC 509
112 ...ThrlYsIleAsnSerSerSerProLeuArgTyrAspArgGlnIleG 127
510 TGGCGCAAGATCTCTAACATGACG.....TTAAGCAACGG 544
127 yGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr.....C 141
545 AAAACTAAGGTTTAACCAAGATGGCTTATTAACCTGTACGCCAACATT 594
141 yGlnValHisPheAspGluGlyLysAlaVal.....TyrLeuLys 154
595 GCTTTCGGCATCATGAACAATCGGGAACGCACTACAGACTATCTTCAG 644
155 LeuAspLeuLeuVal.....AsnGlyValLeuAlaLeuArgCy 167
645 CTGATGGTGTATGCGCTTAACAACGACATCAAAATCCCAAGATCTCATGA 694
167 sleuGluGlnPheSerAlaThrAlaIleAsnSerProGlyProGlnLeuA 184
695 CCGATGACAAAGAGGAGGACGACAAAACTGGTCGGGCAATTCTGATTC 744
184 rGlu.....CysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
745 ACTTTTATTCATMAAAGTTGGGGGATTTTTCACACTCGAGCTGTGTA 794
199 SerLeuArgIleArgThrLeuProThrLalaHisLeuValAlaIleProh 215
795 GAAATTAAGCATTCAGGTGTCACACCTTCCTCGCTGATCCGATCAAGA 844
215 eLeuThrTyrPheGlyLeuPheGlnVal 224
845 TGGCAGCTACTTGGGGCTTCAAAAGTT 872
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/MA1998.DAT.AA413711
seq_documentation_block:
ID AA41371 standard; cDNA; 1630 BP.
XX AA41371:
XX AC
XX DF 08-OCT-1998 (first entry)
XX NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.

```

XX	ANKK: necrosis factor-kappa B; NF-kB: receptor activator; human;
KW	immune response; inflammatory response; toxic shock; sepsis;
KM	RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.
XX	
OS	Mus musculus.
XX	
FH	Key
FT	Location/Qualifiers
CD5	3..887
FT	/tag- ^a
FT	/product- "murine RANKL (ligand for RANK)"
XX	
PN	W09828424-A2.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97MO-US23866.
XX	
PR	14-OCT-1997; 97US-0064671.
PR	23-DEC-1996; 96US-0059978.
PR	07-MAR-1997; 97US-0813509.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Anderson DM, Galibert LJ, Maraskovsky E;
XX	
DR	WPI; 1998-377655/32.
DR	P-PSDB; AAM68292.
PT	New immune receptor activator of necrosis factor-kappa B - useful
PT	for, e.g. developing products for regulating an immune or
PT	inflammatory response, treating toxic shock or sepsis
XX	
SS	Example 7; Pages 55-57; 80pp; English.

This cDNA encodes a murine RANKL, a ligand for the RANK (receptor
CC activator of necrosis factor- κ Bpbp (NF- κ B) polypeptide. RANK is a
CC member of the tumour necrosis factor (TNF) family. Host cells transformed
CC or transfected with an expression vector comprising the RANK encoding
CC nucleic acid can be used to produce recombinant RANK protein. The soluble
CC RANK may be used for inhibiting activation of NF- κ B, by contacting a cell
CC expressing membrane-associated RANK with a soluble RANK which binds to
CC RANK ligand (RANKL). The soluble RANK polypeptide composition may also be
CC used for regulating an immune or inflammatory response. Inhibition of
CC NF- κ B by RANK antagonists may be useful in ameliorating negative effects
CC of an inflammatory response that result from triggering of RANK, e.g. in
CC treating toxic shock or sepsis, graft-versus-host reactions, or acute
CC inflammatory reactions. They can also be used in adjunct therapy for
CC disease characterised by neoplastic cells that express RANK. The products
CC can also be used for detection and drug screening.

SQ Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

XX

alignment_scores:

Quality:	106.50	Length:	276
Ratio:	0.772	Gaps:	13
Percent Similarity:	50.000	Percent Identity:	24.275

alignment_block:
US-09-245-198A-2 x AA4V1371 ..

Align seg 1/1 to: AA4V1371 from: 1 to: 1630

```
3 SerLeuGIyLeuAlaLeuAlaCysLeuGIyLeuLeuLeuValValValSe 19
||||| ||||||| ||||||| ||||| 11
78 TCATGTTCTCGGCCCTTCCTGGGCTTGAGCATGGGCCAGGTGGTCTGCAC 127
19 fLeuGIySerTrpAlaThrLeuSerAlaGln...GlupProSerGIuGLug 35
||||| ::: :||| |::|||:||||:
128 CATCGCTGTGCTGTACTTCGACGGCAGAAAGGAATCCTAACAGA.... 173
35 JuLeuThAlaGIuAspArg..... 41
:::|||||
```


alignment_scores:

Quality: 106.50 Length: 276
Ratio: 0.772 Gaps: 13
Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:

US-09-245-198a-2 x AAD15310 ..

Align seg 1/1 to: AAD15310 from: 1 to: 1630

```

3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValValSe 19
||||| ||||||| ||||||| |||||||
78 TCCAGTTCCTGGCCCTCCCTGGGCTGGGCGACGAGTGGTCTGCAG 127
19 rLeuGlySerTrpAlaThrLeuSerAlaGln..GluProSerGlnGlu 35
||||| ||||||| ||||||| |||||||
128 CATCCCTGTCTCTGCTGCTGCTGAGCGGAGTGCATCTCAACAGA... 173
35 LuLeuThrAlaGluAspArg..... 41
||||| |||||||
174 ....ATATCAGAAAGACAGACACTCACTGCTTTATAGATCCTGAGACTC 218
42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValVa 58
||||| ||||||| ||||||| |||||||
219 CATGAAACGACGATTTGCAGACTCGACTCTGAGAGTGAAGACACACT 268
58 LPro..... 59
|||||
269 ACCTCACTCTCGCAGAGAGATGAACACCTTTCAAGGGGCGCTGCACAGA 318
60 ..PheLeuGlnGlnLeuValArgProArgArg.....SerAlaProlys 73
||||| ||||||| ||||||| |||||||
319 AGGACTCTGCACACATTTGGGGGCCACAGCGCTTCTCAGGACTCTCAGCT 368
74 .....GlyArgLysAlaArgProArgAr 81
||||| |||||||
369 ATGATGGAAGGCTCATGTTGGATGTGGCCCGCAGCGAGCAACCTGAGGC 418
81 GaAlaLeuAlaHisTyrGluValHisProArgProGluGlnAspGlyA 98
||||| |||||||
419 CCAGCCATTTGCACACTCCACCATCAATGCTGCCAGCATCCCATCGGGTT 468
98 IagLInAlaGlyAlaAspGlyThrValSerGlyTrpGluGlu..... 111
||||| ||||||| ||||||| |||||||
469 CCCAT.....AAAGTCACTCTGCTCTCTGTTGACACGATCGAGGC 509
112 ..ThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleG 127
||||| ||||||| ||||||| |||||||
510 TGGGCCAAGATCTTAACATGACG.....TTAAGCAACGG 544
127 YGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr.....C 141
||||| ||||||| ||||||| |||||||
545 AAAAATAAGGGTTAACCAAGATGCTTCTATTACTGTACGCAACATTT 594
141 YSGInValHisPheAspGluGlyLysAlaVal.....TyrLeuLys 154
||||| ||||||| ||||||| |||||||
595 GCTTTCGGCATCATGAACATCGGGAAGCGTACACACATCACTTCGAG 644
155 LeuAspLeuLeuVal.....AsnGlyValIleAlaLeuArgCys 167
||||| ||||||| ||||||| |||||||
645 CTGAGGTGTATGTCTTAACACAGCATCAAAATCCCAAGTTCTCATTA 694
167 sLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuA 184
||||| ||||||| ||||||| |||||||
695 CTTGATGAAGGAGGAGCAGAAAAAAGTGGGCAATTTCTGAATTC 744
184 rGLeu.....CysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
||||| ||||||| ||||||| |||||||
745 ACTTTATTCATTAATGTGGGGATTTTCAAGCTCGGACTGTGTGA 794
199 SerLeuArgGlnLeuArgThrLeuProTrpAlaHisLeuLysAlaAlaProPh 215
||||| ||||||| ||||||| |||||||
795 GAAATTAACATTCAGGTGTCCAACCTTCCCTGATCGGATCCGATCAAGA 844

```

215 eLeuThrTyrPheGlyLeuPheGlnVal 224
||||| ||||||| ||||||| |||||||
845 TCCGACGTACTTGGGGCTTTCAAGTT 872

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAD08714

seq_documentation_block:

ID AAD08714 standard; cDNA; 1630 BP.

AC AAD08714;

DT 04-SEP-2001 (first entry)

DE Murine receptor activator of NF- χ B ligand (RANKL) cDNA.

KW Murine; receptor activator of NF- χ B; RANK; tumour necrosis factor; CD40; TNF receptor-associated factor; TRAF; ligand; immune response; TNF; chromosome 18q22.1; RANKL; chromosome 13q14; transmembrane protein; ss.

KW Mus musculus.

OS Mus musculus.

FX Key Location/Qualifiers

FT CDS

FT /*tag= a

FT /product= "Murine RANKL protein"

FT /note= "CDS does not include start codon"

PN US6242213-B1.

PD 05-JUN-2001.

XX 22-DEC-1997; 97US-0995659.

XX 23-DEC-1996; 96US-0059978.

PR 07-MAR-1997; 97US-0077181.

PR 14-OCT-1997; 97US-0064671.

XX (IMM) IMMUNEX CORP.

XX Anderson DM;

XX WPI: 2001-407216/43.

DR P-P-SDB; AAE04425.

XX New DNA molecules, useful for producing ligands (which are useful for

PT regulating immune response and in screening for inhibitors of NF- χ B

PT receptor activator) of the receptor activator of NF- χ B (RANK)

XX Example 7; Column 57-60; 43pp; English.

XX The present invention relates to receptor activator of NF- χ B (RANK)

XX cDNA. RANK is mapped to chromosome 18q22.1 and its ligand (RANKL) to

XX chromosome 13q14. RANK and RANKL are type 1 and type 2 transmembrane

XX proteins respectively. RANK is a member of the tumour necrosis factor

XX (TNF) superfamily and it closely resembles CD40 in the extracellular

XX region. RANK associates with TNF receptor-associated factor (TRAF) 2

XX and TRAF3. The DNA molecules are useful for producing ligands of RANK.

XX The ligands are useful for regulating immune response and in screening

XX for inhibitors of RANK. The present sequence is murine RANKL cDNA.

SO Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

alignment_scores:

Quality: 106.50 Length: 276
Ratio: 0.772 Gaps: 13
Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:

US-09-245-198a-2 x AAD08714 ..

Align seg 1/1 to: AAD08714 from: 1 to: 1630

3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVal1se 19
|||||
78 TCATGTCTCTGGCCCTCTGGGGGAGGAGTGGCCAGGTGTCTGACG 127
19 rLeuGlySerTrpAlaThrLeuSerAlaGln...GluProSerGlnGluG 35
|||||
128 CATGGCTGTCTCTGCTGACTTTCGAGCGCAGATGATCTCAACAGA... 173
35 LuLeuThrAlaGluAspArg..... 41
174ATATGAGAGACAGCAGCATCTGCTTTATAGATCTGAGACTC 218
42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspVala 58
|||||
219 CATGAAAGCAGATTTGGCAGGAGCTGACCTGTGAGAGTGAACACACT 268
58 LPro..... 59
269 ACCTGACTCTCGAGGAGATGAACAGCCTTTCAGGGGCGCTGCAGA 318
60 ..PheLeuGluGlnLeuValArgProArgArg....SerAlaProLys 73
|||||
319 AGGAACTGCACACATTTGGGGCCACAGCGCTTTCAGGAGCTCCAGCT 368
74GlyArgLysAlaArgProArgArg 81
369 ATGATGGAAGGCTCATGCTGTGATGTGGCCACGAGGAGCAAGCTTGAGGC 418
81 gaLaileAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyA 98
|||||
419 CCAGCCATTTCACACCTCACCATCAATCTGCCAGCATCCATCGGGTT 468
98 laGlnAlaGlyValAspGlyThrValSerGlyTrpGluGlu..... 111
469 CCGCAT.....AAAGTCACCTGTGCTCTTGGTACACGATCGAGGC 509
112 ...ThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleG1 127
|||||
510 TGGGCCAAGATCTCTAACATGAGC.....TTAAGCAAGC 544
127 yGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr.....C 141
|||||
545 AAACATGAAGGTTAACCAAGATGCTTCTATACCTGTACGCCAACATTT 594
141 yGlnAlaHisPheAspGluGlyLysAlaVal.....TyrLeuLys 154
|||||
595 GCTTTGGGCATTCGAAACATCGGGAAGCGTACCTACAGACTATCTTCAG 644
155 LeuAspLeuLeuVal.....AsnGlyValLeuAlaLeuValArgCy 167
|||||
645 CTGATGGTGTATGTCTTAACACAGATCAAAATGCCAAGTTCTCTATA 694
167 sLeuGluGluPheSerAlaThrAlaIleAsnSerProGlyProGlnLeuA 184
|||||
695 CCTGATGAAGAGGAGGAGCAGAAAACCTGTGCGCAATTCGAAATTC 744
184 rGluLeu.....CysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
|||||
745 ACATTTTATTCCTCAATATGTTGGGGATTTTCAAGCTCCAGCTGGGAA 794
199 SerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaIleArgProH 215
|||||
795 GAATTAAGCATTCAGGTGTCCAAACCTTCCTGCTGATCGGATCGAAGA 844
215 eLeuThrTyrPheGlyLeuPheGlnVal 224
|||||
845 TCGCAGCTACTTGGGGCTTCAAAAGTT 872
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD05903
seq_documentation_block:

ID AAD05903 standard; cDNA, 1630 BP.
XX
AC AAD05903;
XX
DT 31-JUL-2001 (first entry)
XX
DE Murine RANKL (receptor activator of NF-kappaB ligand) cDNA.
KW Mouse; receptor activator of NF-kappaB; RANK; nuclear factor-kappaB;
KW NF-kappaB; tumour necrosis factor; TNF; type I transmembrane protein;
KW TNF receptor-associated factor; TRAF; RANK ligand; RANKL; osteopathic;
KW inflammatory reaction; bone resorption; gene therapy; immunomodulator;
KW immune system dysfunction; familial expansile osteolysis; FEO;
KW early onset Paget's disease of bone; Ep; cyostatic; ss.
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 3..887
FT /tag= a
FT /product= "Murine RANKL (receptor activator of
FT NF-kappaB ligand) protein"
XX
XX WO200136637-A1.
XX
XX PD 25-MAY-2001.
XX
XX PF 14-NOV-2000; 2000WO-US31459.
XX
XX PR 17-NOV-1999; 99US-0442029.
XX
XX PA (IMNV) IMMUNEX CORP.
XX
XX PI Anderson DM, Hughes AE;
XX WPI: 2001-329222/34.
XX DR P-PSDB: AAE01992.
XX
XX New DNA encoding a receptor activator of NF-kappaB polypeptide for the
XX treatment of Paget's disease and Familial Expansile Osteolysis (FEO) -
XX
XX Example 7; Page 72-74; 96pp: English.
XX
XX The present invention relates to a novel receptor, referred to as RANK
XX (receptor activator of NF (nuclear factor)-kappaB), a member of TNF
XX CC (tumour necrosis factor) receptor superfamily. RANK is a type I
XX CC transmembrane protein that interacts with TNF receptor-associated
XX CC factors (TRAFs). Triggering of RANK by overexpression or co-expression
XX of RANK and membrane bound RANK ligand (RANKL) results in upregulation
XX of the transcription factor NF-kappaB, a ubiquitous transcription factor
XX CC that is most extensively utilised in cells of the immune system.
XX CC Inhibition of NF-kappaB by RANK antagonists is useful in ameliorating
XX negative effects of inflammatory reactions, and the effects of excess
XX bone resorption. The RANK DNAs, proteins and their analogues are useful
XX for the preparation of pharmaceutical compositions, for infecting target
XX CC cells for use in gene therapy applications in diagnosing diseases
XX CC associated with RANK, and as targets for use in screening assays. They
XX CC may be used in the treatment or diagnosis of immune system dysfunction.
XX CC The present invention also encompasses gene therapy methods to correct
XX CC gene-activating mutations, associated with e.g. familial expansile
XX CC osteolysis (FEO) and early onset Paget's disease of bone (EP). The
XX CC present sequence is a cDNA encoding murine RANKL (murRANKL) protein.
XX
XX Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

alignment_scores: length: 276
Quality: 106.50
Ratio: 0.772 Gaps: 13
Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block: ..
US-09-245-198A-2 x AAD05903

[illegible]

THIS PAGE BLANK (USPTO)

OM of: US-09-245-198a-2 to: GenEmbl:* out-format : pfs
Date: Jun 22, 2002 2:25 PM
About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODE=frame+g2n.model -DEV=x1h
-O=/cgn2.1/USRC.spool/US09245198/runat.21062002.184006.9945/app-query.fasta.1.632
-DB=GenEmbl -QFMT=fastap -SUFFIX=trge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELCP=6.000
-DELEXT=7.000 -START=1 -MATRIX=biom62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09245198.@CGN1.1.10697
-NO_PU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPPY -WAIT -THREADS=1

Search information block:
Query: US-09-245-198a-2
Query length: 225
Database: GenEmbl:*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 2600.680000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb.pr:AF030100	+ 1162.00	1535.91	3.5e-77	1168	AF030100 Mus musculus TWEAK m
gb.pat:AR140407	+ 1020.00	1347.47	1.1e-66	1236	AR140407 Sequence 1 from paten
gb.pr:AF030099	+ 1020.00	1347.02	1.1e-66	1306	AF030099 Homo sapiens TWEAK m
gb.pat:AX201324	+ 1020.00	1346.72	1.2e-66	1353	AX201324 Sequence 3 from Paten
gb.pr:AF055872	+ 1020.00	1346.66	1.2e-66	1368	AF055872 Homo sapiens Ap03/DR2
gb.pat:AX180714	+ 951.00	1258.79	9.4e-62	898	AX180714 Sequence 1 from Paten
gb.pr:BC019047	+ 945.50	1246.46	4.6e-61	1651	BC019047 Homo sapiens, Simlilar
gb.pat:AC069459	+ 602.50	752.45	1.5e-33	203083	AC069459 Mus musculus chrom
gb.ro:AL603707	+ 602.50	751.27	1.8e-33	234182	AL603707 Mouse DNA sequence
gb.htg:AC016876	+ 545.50	682.83	1.1e-29	190358	AC016876 Homo sapiens clone
gb.htg:AC098923	+ 425.50	519.21	1.5e-20	179030	AC098923 Rattus norvegicus c
gb.ov:GGA243435	+ 109.50	144.97	10.34	892	AJ243435 Gallus gallus mRNA for
gb.pr:AL353138	+ 109.50	105.05	1.7e+03	108967	AL353138 Human DNA sequence
gb.da:SC12	+ 108.50	111.60	746.74	42210	AL031124 Streptomyces coelic
gb.htg:AC061974	+ 107.50	119.91	257.31	13243	AE004602 Pseudomonas aerugin
gb.htg:AC060789	+ 107.00	97.65	4.5e+03	178262	AC061974 Homo sapiens chrom
gb.pat:AR156433	+ 106.50	135.99	32.71	1630	AR156433 Sequence 10 from pat
gb.pat:AR164147	+ 106.50	135.99	32.71	1630	AR164147 Sequence 10 from pat
gb.pat:AX147987	+ 106.50	135.99	32.71	1630	AX147987 Sequence 10 from pat
gb.ro:AF019048	+ 106.50	133.41	45.58	2225	AF019048 Mus musculus receptor
gb.pat:AX200995	+ 106.50	128.46	85.92	4034	AX200995 Sequence 625 from Pat
gb.pat:AX267730	+ 106.50	128.46	85.92	4034	AX267730 Sequence 704 from Pat
gb.pat:AX200993	+ 106.50	126.86	105.57	4894	AX200993 Sequence 623 from Pat
gb.pat:AX267728	+ 106.50	126.86	105.57	4894	AX267728 Sequence 702 from Pat
gb.pat:AX200996	+ 106.50	123.91	154.01	6976	AX200996 Sequence 626 from Pat
gb.pat:AX267731	+ 106.50	123.91	154.01	6976	AX267731 Sequence 705 from Pat
gb.htg:AC096533	+ 106.50	97.99	4.3e+03	157988	AC096533 Homo sapiens chrom
gb.htg:AC090975	+ 106.00	96.46	5.2e+03	175347	AC090975 Mus musculus chrom
gb.htg:AC073779	+ 106.00	95.06	6.2e+03	207585	AC073779 Mus musculus chrom
gb.htg:AC073801	+ 106.00	94.50	6.7e+03	222037	AC073801 Mus musculus clone
gb.pr:AY070219	+ 105.50	126.39	112.01	4412	AY070219 Homo sapiens lympho
gb.pr:HUMLYTOXB	+ 105.50	123.43	163.87	6305	L11016 Homo sapiens lympho
gb.htg:ALA51008	+ 105.50	104.39	1.9e+03	62337	AL451008 Homo sapiens chrom
gb.htg:AL360009	+ 105.50	104.39	1.9e+03	62337	AL360009 Homo sapiens chrom
gb.pr:HSY14768	+ 105.50	102.14	2.5e+03	81800	Y14768 Homo sapiens DNA, cosm
gb.pr:AP000505	+ 105.50	100.47	3.1e+03	110000	AP000505 Homo sapiens genom
gb.htg:LMICRHL6_09	+ 105.50	99.67	3.4e+03	110000	Continuation (10 of 11) of L
gb.htg:AL662801	+ 105.50	97.71	4.4e+03	179359	AL662801 Homo sapiens chrom
gb.htg:AL662847	+ 105.50	96.03	5.5e+03	170629	AL662847 Homo sapiens chrom

gb.pr:DJ0201G24	+ 105.50	95.37	6.0e+03	184666	AF129756 Homo sapiens MSN
gb.pl:AP002482	+ 105.50	95.23	6.1e+03	187835	AP002482 Oryza sativa gen
gb.ba:AE005079	+ 104.00	115.56	449.43	12797	AE005079 Halobacterium sp.
gb.htg:AC061972	- 104.00	94.35	6.8e+03	164425	AC061972 Homo sapiens clo
gb.pr:AC079384	+ 103.50	92.91	8.2e+03	180555	AC079384 Homo sapiens 12q

seq_name: gb.ro:AF030100

seq documentation block:	1168 bp	mRNA	linear	ROD 20-DEC-1997
LOCUS	AF030100			
DEFINITION	Mus musculus TWEAK mRNA, partial cds.			
ACCESSION	AF030100			
VERSION	AF030100.1	GI:2707220		
KEYWORDS	house mouse.			
SOURCE	Mus musculus.			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 1168)			
AUTHORS	Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L.			
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis			
JOURNAL	J. Biol. Chem. 272 (51), 32401-32410 (1997)			
MEDLINE	98070415			
REFERENCE	2 (bases 1 to 1168)			
AUTHORS	Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.			
TITLE	Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA			
JOURNAL				

FEATURES

source

1..1168

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_type="peritoneal macrophages"

<1..679

/note="ligand in the TNF family; secreted protein"

/codon_start=2

/product="TWEAK"

/protein_id="AAC53517.1"

/db_xref="GI:2707221"

/translation="VLSIGLALCLGLLVVSLGSMATLSAQPSQELTADREPP

PELNPOTEEQDVPFLQVLRPSAPKRRAPRAIAHVEVPRPODGAQGV

DGVNSMEETKINSSPLRYDRDGETVTRADLYLVCOVHDEKAVYIKDLVY

GVIALRCLEFSATSAASPPQRLCOVSSLLPLRSGSLRITRLMAHLKAPFLTY

FGLFQYH"

BASE COUNT 242 a 360 c 298 g 268 t

ORIGIN

alignment_scores:

Quality: 1162.00 Length: 225

Ratio: 5.164 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-2 x AF030100

Align seg 1/1 to: AF030100 from: 1 to: 1168

1 ValLeuSerLeuGlyLeuAlaLeuAlaLeuAlaLeuLeuValVa 17

|||||

2 GTCGCGACCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCTGGT 51

|||||

17 lValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGln 34

|||||

52 GGTGCGCTGGGAGACTGGCAACGCTGTCGCCAGAGACCTTCAGG 101

|||||

34 lValLeuThrAlaGlnAspArgArgGluProProGluLeuAspProGln 50

|||||

102 AGGACTACAGCAGAGACCGCGGAGCCCGCTGGAAGTATCCAG 151

|||||

51 ThrGluGluSerGlnAspValValProPheLeuGluGlnLeuValArgPr 67

```

|||||
152 ACAGAGAAAGCCAGATGTGTACTTCTTGGAACAACACTAGTCCGGCC 201
67 oarfgaSerAlaProLySGlyValArgValaArgProArGAla11a 84
|||||
202 TCGAAGAAGTGTCTTAAGAGCCGGAAGGCGCGCTCCGCGAGCTATTG 251
84 la1a1a1sTyrgLuvAla1sProArGProGlyGlnAspGlyAlaGlna1a 100
|||||
252 CAGCCATTATGAGTTGATCTCGGCCAGAGACAGATGAGACACAAGCA 301
101 GlyValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerse 117
|||||
302 GGGTGGATGCGACAGTGAAGTGGTGGGAAGACCAAAATCAACAGCTC 351
117 rSerProLeuArgTyRAspArgGlnIleGlyGluPheThrVal11aArgA 134
|||||
352 CAGCCCTGCGGCTACGACCGCCAGATTGGGAAATTACAGTATCAGGG 401
134 laGlyLeuTyRtyrLeuTyRcysGlnVala1sPheAspGlyGlyAla 150
|||||
402 CTGGGCTCTACTACTGTACTGTGACAGTGTGATGAGGAAAGGCT 451
151 ValTyRLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgcy 167
|||||
452 GTCTACTGTAAGCTGACTGTGTAACGGTGTGCTGCGCCCTGCGCTG 501
167 slauGluGluPheSerAlaThrAla1a1aSerSerProGlyProGlyLeu 184
|||||
502 CCGGAAGAATTCTCAGCCACAGCAGCAGACCTCTCGGCGCCAGCTCC 551
184 rgluGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
|||||
552 GTTGTGCCAGGTGTGCGGCTGTGCGCTGCGCGCAGGAGTCTTCCCTT 601
201 ArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuTh 217
|||||
602 CGGATCCGACCCCTCGGCTCATTTAAGGTCGCCCTTCTCTAAC 651
217 rTyRPhgGlyLeuPheGlnValHis 225
|||||
652 CTACTTGGACTCTTTCAGACTTAC 676

```

seq_name: gb_pat:ARI40407

seq_documentation_block:

LOCUS ARI40407 1236 bp DNA Linear PAT 16-JUN-2001
 DEFINITION Sequence 1 from patent US 6207642.
 ACCESSION ARI40407
 VERSION ARI40407.1 GI:14482903

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1236)

AUTHORS Wiley,S.R.

TITLE Member of the TNF family useful for treatment and diagnosis of

JOURNAL Patent: US 6207642-A 1 27-MAR-2001;

FEATURES Location/Qualifiers

BASE COUNT 225 a 416 c 358 g 237 t

ORIGIN 1.1236 /organism="unknown"

alignment_scores:

Quality: 1020.00 Length: 224

Ratio: 4.744 Gaps: 0

Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198a-2 x ARI40407

Align seg 1/1 to: ARI40407 from: 1 to: 1236

```

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValVala 18
|||||
76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 125
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
126 CATTGTTGGGAGCCGGCATCGCTGTCCGCCAGAGAGCTGCCAGAGAG 175
35 lLeuThrAlaGluAspArgArgGluProProGlyLeuAsnProGlnThr 51
|||||
176 AGCTGTGGCAGAGAGAGACAGACACCGCTCGGAATGATCCCGAGAA 225
52 GluGlySerGlnAspValAlaProPheLeuGluGlnLeuValaArgPro 68
|||||
226 GAGAAAGCCAGATCTCGGCTTCTCTGAAACCGACTGCTGGGCTGG 275
68 garSerAlaProLySGlyValArgValaArgProArGAla11aAla 85
|||||
276 AAGAGTGACCTAAGCGCGGAAGACAGCGGCTCGAAGAGCATCGAG 325
85 laHisTyrgLuvAla1sProArGProGlyGlnAspGlyAlaGlna1aGly 101
|||||
326 CCATTATGAAATTCATCCAGCAGCTGAGACAGAGACAGCGCGGAGAT 375
102 ValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerse 118
|||||
376 GTGGAGCGGACAGTGAAGTGGTGGGAGAAAGCCAGATCAACAGCTCC 425
118 rProLeuArgTyRAspArgGlnIleGlyGluPheThrVal11aArgAlaG 135
|||||
426 CCTGTGCGCTACACCGCAGATCGGGAGTTATAGTACACCGGGCTG 475
135 lLeuTyRtyrLeuTyRcysGlnVala1sPheAspGlyGlyAlaVala 151
|||||
476 GGCTCTACTACCTGTACTGTGACAGTGCACCTTGTATGAGGAGGCTG 525
152 TyRLeuLysLeuAspLeuVala1sGlyValLeuAlaLeuArgCysLe 168
|||||
526 TACTGAGAGCTGACTGTGCTGGTGTGATGTGTGCTGCGCTGGCTG 575
168 uGluGluPheSerAlaThrAla1a1aSerSerProGlyProGlyLeuArg 185
|||||
576 GAGGAATTTCTCAGCCATCGCGGAGTTCCTCGGCGCCAGCTCCGCC 625
185 euGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
626 TGTGCCAGGTGTGCGGCTGTGCGCTGCGGCGCAGAGTCTCTGCGG 675
202 lLeArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuTh 218
|||||
676 ATCCGACCCCTCCCTGGGCGCCATCTCAAGCTGCGCCCTCTCTAC 725
218 rPhgGlyLeuPheGlnValHis 225
|||||
726 CTTCGAGCTCTTCCAGGTTTAC 747

```

seq_name: gb_pr:AF030099

seq_documentation_block:

LOCUS AF030099 1306 bp mRNA Linear PRI 20-DEC-1997
 DEFINITION Homo sapiens TWEAK mRNA, complete cds.
 ACCESSION AF030099
 VERSION AF030099.1 GI:2707218

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1306)
 Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,

AUTHORS

FEATURES	source	CDS	BASE COUNT	ORIGIN
JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL	Hession,C., Garcia,I. and Browning,J.L. TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis J. Biol. Chem. 272 (51), 32401-32410 (1997) 2 (bases 1 to 1306) Bourdon,P., Hession,C., Tizard,R. and Browning,J. Direct Submission Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA		247 a	
	Location/Qualifiers 1..1306 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17p13" /tissue_type="tonsil" 18..767 /note="ligand in the TNF family; secreted protein; start codon not verified experimentally" /codon_start=1 /product="TWEAK" /protein_id="AAC51923.1" /db_xref="GI:2707219" /translation="MAARSRGRRGRGEGPTALIVPLALGLALACLGILAAVSLGSRSLAEOEPAOELEEDODPSELNPTEESODAPAFINLRVRRSAPKGRTRARRAIAAHVEYHPRGDDGAAGVDGVSMEERKRISSSPINRQIGETIVRALLYLICQVHFDEGKAVYILKDLIDLVGYALAKLEEFSTAASSLGAPQRLCOVSGLLALRPSPSLRIRLIPWMLHAAPLIFYGLFQVH"		434 c	368 g
			257 t	

```

alignment_scores:
    Quality: 1020.00      Length: 224
    Ratio: 4.744          Gaps: 0
    Percent Similarity: 95.982      Percent Identity: 88.839

alignment_block:
US-09-245-198A-2 x AF030099 ..

Align seg 1/1 to: AF030099 from: 1 to: 1306

2 LeuSerGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVala 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
93 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGT 142
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnClnProSerGlnG 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
143 CAGTTTGGGGAGCCGGGGATCGCTGTCCGCCAGGAGCCTGCCCAAGAG 192
35 lLeuThrAlaGlnAspArgArgGlnProProGlnLeuAsnProGlnThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
193 AGCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
52 GlnGlnSerGlnAspValValProPheLeuGlnLeuValAlaArgProAr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
243 GAAGAAAGCCAGAGATCTCGCCCTTTCTGAACCACTAGTTCCGGCTCG 292
68 gatSerAlaProGlyGlyArgGlyAlaArgProArgArgAlaIleAla 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293 CAGAGATGCACCTTAAGGCCGGAAACACAGGGCTGAAAGAGGATGCAG 342
85 lAla1TrpGlnValAla1SerProArgProGlyGlnAspGlyValaGlnAlaGly 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
343 CCCATTATGAAGTTCATCCAGAGACTGGAGACAGAGAGAGAGAGAGAG 392
102 ValAspGlyThrValSerGlyTrpGlnGlnThrLysIleAsnSerSerSe 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
393 GTGGACGCGAGACAGTACGTGGCTGGAGGAAGACCAATCAACAGCTCCAG 442
118 rProLeuArgTrpArgPargGlnIleGlyGlnPheThrValIleArgAlaG 135
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

443 CCCCTCGGGGTACACCGGCAGATCGGGGGAGTTTATAGTCACCCGGGCTG 492
135 lYleuTyrTrIleuTyrCysGlnValHisPheaspcluglylsalaval 151
|||||
493 GGCTCTACTACCTGTACTGTCAAGTCACTTTGATGAGGGGAGGCTGTC 542
152 TyrLeuLysLeuaspLeuValasnGlyValLeuAlaLeuargCysLe 168
|||||
543 TACTCTAACCTGGACTTGCTGTGATGTGTCTGGCCCTGCGCTGCTCT 592
168 uGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeuArg 185
|||||
593 GAGGAGATTCCTCAGCCACCTCGGCCAGTTCCTCCGCGGCCCACTCCGCC 642
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
643 TCTGCCAGGTGTCTGTGGCTGTGGCCCTCGCGGCCAGGATCTCTCGGG 692
202 lIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 218
|||||
693 ATCCGACACCTCCCTCGGGGCCCATCTCAAGGCGTCCCTCTCACCTA 742
218 rPheGlyLeuPheGlnValHis 225
|||||
743 CTTCGGACTCTTCAGGTTCCAC 764

seq_name: gb_pat:AX201324

seq_documentation_block:
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1353)
AUTHORS
Hillan,K.J., Marsters,S.A., Pan,J., Pittl,R.M., Roy,M.A., Smith,V.,
Stone,D.M., Watanabe,C.K. and Wood,W.I.
Compositions and methods for the treatment of tumour
Patent: WO 0153486-A 3 26-JUL-2001;
Genentech, Inc. (US)
FEATURES
location/Qualifiers
source 1..1353
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:
US-09-245-198A-2 x AX201324 ..
Align seg 1/1 to: AX201324 from: 1 to: 1353
2 leuSerleuGlyleuAlaLeuAlaCysleuGlyleuLeuValVal 18
|||||
133 CTGGGGCTGGGCTCGGCGGCTCTCTCGGCTCTCTCTCTCTCTCTCTCTCTG 182
18 lserleuGlySerTrpAlaThrleuSerAlaGlnGluProSerGlnGlu 35
|||||
183 CAGTTTGGGAGCGCGGCACTGCTGTCCGCCAGAGAGCTGCCAGAGG 232
|||||
18 luleuThrAlaGluaspArgGluProProGluLeuasnProGlnThr 51
|||||

```

```
233 AGCTGTCGACAGAGACAGACCGACCGCTCGAATCGATCCCGACAGA 282
52 GUGUGJUSerGlnAspVal1ProPheLeuGluGlnLeuValArProAr 68
283 GAAGAAAGCCAGATCCTGGCTTCTCTGAAACGACTAGTGGGCTCG 332
68 GARSerAlaProLyGlyArGlyAlaArProArGArGAlaIleAla 85
333 CAGAAATGCACCTAAAGCGCGGAAACACGGGCTCGAAGGCGATCGCAG 382
85 laH1sTyrGlnVal1HisProArGProGlyGlnAspGlyAlaGlnAlaGly 101
383 CCCATTATGAACTTCATCCACGACTGACAGACGACGACGACGACG 432
102 ValAspGlyThrValSerGlyTrpGluGluThrVal1LeuAsnSerSer 118
433 GUGAGCGGACAGACTGAGTGGCTGGAGGAACCAATCAACGCTCCAG 482
118 rProLeuArGlyTyrAspArgGlnIleGlyGluPheThrVal1IleArGAlaG 135
483 CCTCTGCCCTACACCGCCAGATCGGGAGTTATAGTCACCCGGGCTG 532
135 lYleuTyrTrLeuTyrCysGlnVal1HisPheAspGlyGlyVal1Val 151
533 GGCTCTACTACTGTAAGTGTCAAGTGTGATGAGGGAGGCTGTG 582
152 TyrLeuLysLeuAspLeuVal1AsnGlyVal1LeuAlaLeuArGcysLe 168
583 TACCTGAAGCTGACTTGGCTGGTGGATGGTGTCTGGCCCTGGCGCTG 632
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgL 185
633 GGAGGAATTCCTCAGCCACTGCGCGGAGTTCCCTCGGCGCCACCTCGCG 682
185 eucysGlnVal1SerGlyLeuLeuProLeuArProGlySerSerLeuArg 201
683 TCTGCCAAGTGTCTGGCTGTGGCTGGCTGGCGGACGAGTCTCTCGCG 732
202 lIeArGThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTy 218
733 ATCCGCAACCTCCCTGGGCGCCATCTCAAGGCTGCCCTCTCTCACTCA 782
218 rPheGlyLeuPheGlnVal1His 225
783 CTTCGAGACTTCTCCAGATTAC 804

seq_name: gb_pr:AF055872
documentatation_block:
DEFINITION Homo sapiens APO3/DR3 ligand (APO3L) mRNA, complete cds.
ACCESSION AF055872
VERSION AF055872.1 GI:3108230
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1368)
Masters, S.A., Sheridan, J.P., Pitti, R.M., Brush, J., Goddard, A. and
Ashkenazi, A.
IDENTIFICATION OF A LIGAND FOR THE DEATH-DOMAIN-CONTAINING RECEPTOR
APO3
JOURNAL Curr. Biol. 8 (9), 525-528 (1998)
MEDLINE 98228355
REFERENCE
2 (bases 1 to 1368)
Masters, S.A., Sheridan, J.P., Pitti, R.M., Brush, J., Goddard, A. and
Ashkenazi, A.
TITLE
Direct Submission
JOURNAL Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,
South San Francisco, CA 94080, USA
FEATURES
source location/Qualifiers
1..1368
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
1..1368
/gene="APO3L"
58..807
/gene="APO3L"
/function="binds to the death-domain-containing receptor
APO3/DR3"
/note="TNF homolog"
/codon_start=1
/product="APO3/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARRSQRRRGREPEPTALLVPLAGLGLALGLLAVSL
GSRASLSQEPDQAEIVAEEDPDSELPDTEESODAPFLNLRPRRSAPKGRK
ARRAIAAHYEVHPRPGDDGADGADGVDSYSGEERINSSFLRNQIGETIVRAGL
VYLQVHFDEGKAVYLLKLDLVDGVALARLCEEFSAITAASSLGPDLRLQVSGLLAL
RPGSLRLRLTPMAHLKNAFLTYGTFQVH"
BASE COUNT 272 a 443 c 389 g 264 t
ORIGIN
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839
alignment_block:
US-09-245-198a-2 x AF055872 ..
Align seg 1/1 to: AF055872 from: 1 to: 1368
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuVal1Val1 18
133 CTGGGCTCGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 182
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
183 CAGTTTGGGGAGCGCGGACGCTGTCTGCCCGCAGAGCGTGGCCAGAG 232
35 lLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGlnThr 51
233 AGCTGTCGCGAGAGAGACCAAGACCGCTCGGAGACTGATCCCGACAGA 282
52 GUGUGJUSerGlnAspVal1ProPheLeuGluGlnLeuValArProAr 68
283 GAAGAAAGCCAGATCCTGGCTTCTCTGAAACGACTAGTGGGCTCG 332
68 GARSerAlaProLyGlyArGlyAlaArProArGArGAlaIleAla 85
333 CAGAAATGCACCTAAAGCGCGGAAACACGGGCTCGAAGGCGATCGCAG 382
85 laH1sTyrGlnVal1HisProArGProGlyGlnAspGlyAlaGlnAlaGly 101
383 CCCATTATGAACTTCATCCACGACTGACAGACGACGACGACGACG 432
102 ValAspGlyThrValSerGlyTrpGluGluThrVal1LeuAsnSerSer 118
433 GUGAGCGGACAGACTGAGTGGCTGGAGGAACCAATCAACGCTCCAG 482
118 rProLeuArGlyTyrAspArgGlnIleGlyGluPheThrVal1IleArGAlaG 135
483 CCTCTGCCCTACACCGCCAGATCGGGAGTTATAGTCACCCGGGCTG 532
135 lYleuTyrTrLeuTyrCysGlnVal1HisPheAspGlyGlyVal1Val 151
533 GGCTCTACTACTGTAAGTGTCAAGTGTGATGAGGGAGGCTGTG 582
152 TyrLeuLysLeuAspLeuVal1AsnGlyVal1LeuAlaLeuArGcysLe 168
583 TACCTGAAGCTGACTTGGCTGGTGGATGGTGTCTGGCCCTGGCGCTG 632
```

168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgL 185
|||||
633 GGAGGAATTTCTAGCCACTGCGCGCAGATTCCCTCGGGCCCCAGCTCCGCC 682
717
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
683 TCTGCAGAGTGTCTGGCTGTGGCCCTGGCGCAGAGGTCTCTCCCTGGCG 732
722
202 IleArgThrLeuProThrAlaHisLeuValAlaProPheLeuThrTy 218
|||||
733 ATCCGACCCCTCCCTGGCCCTCAAGCTCCCTCCCTCAACCTA 782
218 rPheGlyLeuPheGlnValHis 225
|||||
783 CTTCGAGCTCTTCAGGTTCCAC 804
seq_name: gb_pat:AX180714
seq_documentation_block:
ACCESSION AX180714 898 bp DNA linear PAT 06-AUG-2001
VERSION AX180714 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 898)
AUTHORS Willey,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
IMMUNEX CORPORATION (US)
FEATURES
source location/Qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
CDS
translation="MATGSRSTLLAFGLCLPMLDEGATSSDRMKOIEKIEILS
KIYHINETAIRIKLIGERTSSASASQEPQELVAEDDPSFINTEES
QDPAPFLNLYRRRSAPKRTKRAKALAHAEVHPQDGAQGVDTVSAMEBA
RINSSSPLRNRQIGERTVRAGLYLCQVHDEGAVALKLDLVDGLARCLLE
FSATPAASSLQPOLRLCOVSGLLALRPSSLRIRTLPAHLKAAFLTYEGIQVH"
BASE COUNT 187 a 266 c 267 g 178 t
RIGIN
alignment_scores:
Quality: 951.00 Length: 207
Ratio: 4.803 Gaps: 0
Percent Similarity: 95.652 Percent Identity: 88.889
alignment_block:
US-09-245-198a-2 x AX180714 ..
Align seg 1/1 to: AX180714 from: 1 to: 898
19 SerLeuGlySerThrPalaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
250 AGTTTGGGAGCGGGGCGATCGCTGCCGCCAGGACCTCGCCAGAGAGA 299
35 uLeuThrAlaGluAspArgGluProGluLeuAsnProGlnThrG 52
|||||
300 GCTGTGCGAGAGAGACAGACAGACCCGCGAATGATCCCGACAGAG 349
52 uGluSerGlnAspValProPheLeuGlnGluLeuValArgProArg 68
|||||
350 AAGAAAGCCAGAGATCTGCGCTTCGACGAGCTAGTTCGGCTCGC 399

69 ArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAla 85
|||||
400 AGAAGTGCACCTAAAGGCCGGAACACGGGCTCGAAGGCGCATGCGAC 449
85 aHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAglY 102
|||||
450 CCATTATGAAATTCATCCACGACCTGAGACAGAGGAGGAGGAGGAG 499
102 aLaspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSer 118
|||||
500 TGGAGGGGACATGATGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 549
119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAl 135
|||||
550 CCTGTGCGTACACCGCCAGATCGGGAGATTATAGTCACCCGCGTGG 599
135 yLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 152
|||||
600 GCTTACTACCTGTAAGTGTGACAGTGTGACCTTTGATGAGGGAGGCTGTCT 649
152 yLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeu 168
|||||
650 ACCTGAAGCTGACCTTGTGCTGGATGATGATGCTGCTGCGCTGCGCTG 699
169 GluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArg 185
|||||
700 GAGGAATTTCTAGCCACTGCGCGCAGTTCCCTCGGGCCCCAGCTCGGCT 749
185 uCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 202
|||||
750 CTGCAGAGGTGCTGGCTGTGGCCCTGGCGCAGAGGTCTCTCCGCGGA 799
202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTy 218
|||||
800 TCCGACACCTCCCGCGGCGCATCTCAAGGCGCCCTCTCTCCTACCTAC 849
219 PheGlyLeuPheGlnValHis 225
|||||
850 TTTCGAGCTCTTCAGGTTCCAC 870
seq_name: gb_pr:BC019047
seq_documentation_block:
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgcenhgr1.nih.gov

Shvchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Boufard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Kalins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C., McBowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 30 Row: P Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.

FEATURES

source

Location/Qualifiers

1..1651

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:20669 IMAGE:4766071"

/tissue_type="Primary B-Cells from Tonsils"

/clone_lib="NIH_MGC_48"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

106..510

/codon_start=1

/product="Similar to tumor necrosis factor (ligand)

superfamily, member 12"

/protein_id="AAH19047.1"

/db_xref="GI:17512139"

/translation="MAARSRORRRGREGPTALLPLALGIALACTIGLLAVSL

GSRASISNOPEAEIVAEEDPSSEINPOTFEESDAPFLNRLVPRKSRPKRKR

ARRAIAHYEHPRPGDGAQADGGYTTCLRP"

BASE COUNT

344 a 517 c 481 g 309 t

ORIGIN

alignment_scores:

Quality:	945.50	Length:	304
Ratio:	4.418	Gaps:	2
Percent Similarity:	70.395	Percent Identity:	65.132

alignment_block:

US-09-245-198a-2 x BC019047 ..

Align seg 1/1 to: BC019047 from: 1 to: 1651

```
2 LeuserLeuglyLeuAlaLeuAlaCysLeuglyLeuLeuLeuValVal1a 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 CTGGGCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 230

18 LserLeuglySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 CAGTTTGGGAGCGCGGCGATCGCTGCCGCCAGAGCGCCAGAGAGG 280

35 LuLeuThrAlaLysAspArgArgGluProProGluLeuAsnProGlnThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
281 AGCTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330

52 GluGluSerGlnAspValValProPheLeuGluGlnLeuValArgProAr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
331 GAAGAAAGCAGAGATCGTGGCCTTTCCTGAACGACTACTTGGCCTCG 380

68 garGserAlaProLysGlyArgLysAlaArgProArgArgAlaIleAla 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 CAGAGAGTGCACCTAAAGGCCGGAAGACACGCGCTCGAAGAGCAGTCG 430

85 lalIstYrGluValHisProArgProGlyGlnAspGlyAlaGln..Ala. 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
431 CCCATTATGAGTTCATCCAGACTGGACAGAGAGAGAGAGAGAGAGAGAT 480

100 ..... 100

481 GGAGGTTACACACTTGTCTGAGGCCATGAGTACTAAGTGTGTGAGCCCA 530
```

```
100 ..... 100
531 AGATTGAACCCAGCTAGAATGTGCTGTACTGACATGATGTCTGCANG 580
100 ..... 100
581 AAGCGAGGTGGCTGTGCAGGGGTGAGGGGTCCATGCAGGGGCCACAT 630
100 ..... 100
631 CCAAAAGGGAGAGGAGAGATTCCAGAAAGAGAGAGACACATCTCCC 680
101 ..... GlyValAspGlyTh 105
681 ACCATTACAGAGGGTCAAAAGAGAAACAGACACCGCTGGACCGGAG 730
105 rValserGlyTrpGluGluThrLysIleAsnserSerProLeuArgP 122
731 AGTGAAGTGGCTGGAGAGAGCCAGAAATCAACAGCTCCAGCCCTGTG 780
122 yAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuYrYr 138
781 ACAAGCCGAGATCGGGAGATTATAGTCACCGGGCTGGGCTCTACTAC 830
139 LeuYrCysGlnValHisPheAspGluGlyLysAlaValYrLeuLys 155
831 CTGTACTGTGACGGTCACTTGTATGAGGGGAAGGCTGTCTACTGAAG 880
155 uAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGluGluPh 172
881 GGACTGTGGTGTGATGTGTGCTGGCCCTCGCCTCCCTGGAGAAATCT 930
172 eAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnVal 188
931 CAGCCACTGGGGGAGATTCCTCGGGCCAGCTCCGCTCGCAGAGTG 980
189 SerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgTh 205
981 TCTGGGCTGTGGCCCTCGGGCCAGGGGTCTCCCTCGGATCCGACCT 1030
205 uProTrpAlaHisLeuLysAlaAlaProPheLeuThrYrPheGlyLeu 222
1031 CCCCTGGGCCCATCTCAAGGCTGCCCTTCTCCTCAGCTTCTGAGCTCT 1080
222 heGlnValHis 225
1081 TCCAGGTTTCAC 1091

seq_name: gb_hc9:AC069459
seq_documentation_block:
LOCUS AC069459 203083 bp DNA linear HTG 27-JUN-2001
DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC069459 GI:14547768
VERSION AC069459.23
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 203083)
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
Dederich, D., Thomas, S., Okwuonu, G., Carllock, C., Garner, T.,
Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,
Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,
Cox, C., Davis, C., Delgado, O., Ding, T., Dugan-Kocha, S.,
Fernandez, C., Ferraguto, D., Forcum-Tansley, J., Gill, R.,
Gortrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hoques, M.,
Hosack, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,
Kovar, C., Liu, J., Liu, W., Louisgeed, H., Lozada, R.J., Martin, R.,
Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,
Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogun, M., Parish, B.,
```


This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6 This sequence is the entire insert of clone RP23-422L16.

FEATURES

Source
1. 234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPI-23"
complement(84050..84131)
/note="Sequence from uni-directional primer reads and dgrp big dye terminator reads only."
misc_feature
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

alignment_scores:
Quality: 602.50 Length: 180
Ratio: 4.782 Gaps: 2
Percent Similarity: 70.000 Percent Identity: 68.889

alignment_block:

US-09-245-198a-2 x AL603707/rev ..

Align seg 1/1 to reverse of: AL603707 from: 1 to: 234182

```

100 ALGlyAlaspLythrhValSerGlyTrpGluThrIleuLeu 116
    ::::::::::::::::::::::::::::::::::::::::::::
66609 TCAGGCTGGATGGAGACGTGAGTGGGGAAGACCAAAATCAACAG 66560
116 rSerSerProLeuArGTYrAspArgInIleGlyGluPheThrValIleA 133
    ::::::::::::::::::::::::::::::::::::::::::::
66559 CTCAGGCCCTCTGCGCTACGACGCCCGCATTTGGGAATTACAGTCATCA 66510
133 rGAlaGlyLeuTYrTYrLeuTYrCys_Gln..... 142
    ::::::::::::::::::::::::::::::::::::::::::::
66509 GGCGTGGGCTCTACTACTCTACTGTCAGTAAAGCCCTGCGCTGCCTCAG 66460
142 ..... 142
66459 GGGCAGACGAAGGCTAAGGGAGAGGAGACTGGCAAGAAAGTGGTGGA 66410
142 ..... 142
66409 GAGGAGACTCTGCGACTCATGAAGAAGATGCCCTGCGCTTCGATGAGG 66360
143 .....ValHisPheAs 146
        ::::::::::::::
66359 AGGCAGAGCTCTGGATTGCTTCCTCTCTCTGTTCAGAGTGCACTTGA 66310
146 pGluGlyLyAlaValTYrLeuLysIleuAspLeuValAsnGlyValL 163
66309 TGAAGGGAAGAGCTGTCTACTGGAAGCTGAGCTGTGCTGAACGGTGTGC 66260
163 euAlaLeuArGcYsLeuGluGluPheSerAlaThrAlaIleSerSerPro 179
    ::::::::::::::::::::::::::::::::::::::::::::
66259 TGGCCCTGGCGCTGCTGGAAGAAATTCACGCCACAGCAGCAAGCTCTCT 66210

```

```

180 GLYProGlnLeuArGcYsGlnValSerGlyLeuLeuProLeuArGp 196
    ::::::::::::::::::::::::::::::::::::::::::::
66209 GGGCCCCAGCTCCGTTGTCAGAGTGCTGCGCTGTGCCCTCGCGGCC 66160
196 ogYserSerLeuArGIlleArGThrLeuProTrpAlaHisLeuLYAlaA 213
    ::::::::::::::::::::::::::::::::::::::::::::
66159 AGGGTCTCCCTTGGATCCGACCCCTCCCTGCGCTCACTTAAAGCTG 66110
213 lArProPheLeuThrTYrPheGlyLeuPheGlnValHis 225
    ::::::::::::::::::::::::::::::::::::::::::::
66109 CCCCCTTCCTCAACTACTTGGACTCTTCAAGTTCCAC 66072

```

seq_name: gp_hcg:AC016876

seq_documentation_block:
LOCUS AC016876 190358 bp DNA linear HTG 03-SEP-2001
DEFINITION Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered pieces.
ACCESSION AC016876 GI:15421989
VERSION AC016876.4
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 190358)
JOURNAL Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE
AUTHORS 2 (bases 1 to 190358)
Unpublished Homo sapiens chromosome, clone RP11-186B7

Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2001 this sequence version replaced gi:13431059.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 13849
Center clone name: 186.B.7
----- Summary Statistics
Sequencing vector: M13; W77815; 3% of reads
Sequencing method: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181928 bases at least Q40
Consensus quality: 185529 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 188858; sum-of-ctrls
Quality coverage: 8.1 in Q20 bases; agarose-fp
Quality coverage: 7.2 in Q20 bas.

* NOTE: This is a 'working draft' sequence. It currently consists of 16 conligs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 20703: contig of 20703 bp in length
* 20704 20803: gap of 100 bp
* 20804 21935: contig of 1132 bp in length
* 21936 22035: gap of 100 bp
* 22036 24202: contig of 2167 bp in length
* 24203 24302: gap of 100 bp
* 24303 28199: contig of 3897 bp in length
* 28200 28299: gap of 100 bp
* 28300 35021: contig of 6722 bp in length
* 35022 35121: gap of 100 bp
* 35122 40745: contig of 5624 bp in length
* 40746 40845: gap of 100 bp
* 40846 47136: contig of 6291 bp in length
* 47137 47236: gap of 100 bp
* 47237 58980: contig of 11744 bp in length
* 58981 59080: gap of 100 bp
* 59081 68164: contig of 9084 bp in length
* 68165 68264: gap of 100 bp
* 68265 94625: contig of 26361 bp in length
* 94626 94725: gap of 100 bp
* 94726 102564: contig of 7839 bp in length
* 102565 102664: gap of 100 bp
* 102665 116764: contig of 14100 bp in length
* 116765 116864: gap of 100 bp
* 116865 136408: contig of 19544 bp in length
* 136409 136508: gap of 100 bp
* 136509 156380: contig of 19872 bp in length
* 156381 156480: gap of 100 bp
* 156481 180631: contig of 24151 bp in length
* 180632 180731: gap of 100 bp
* 180732 190358: contig of 9627 bp in length.
Location/Qualifiers
1. 190358

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPC1-11 Human Male BAC"
1. 20703
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
20804..21935
/note="assembly_fragment"
misc_feature
22036..24202
/note="assembly_fragment"
misc_feature
24303..28199
/note="assembly_fragment"
misc_feature
28300..35021
/note="assembly_fragment"
misc_feature
35122..40745
/note="assembly_fragment"
misc_feature
40846..47136
/note="assembly_fragment"
misc_feature
47237..58980
/note="assembly_fragment"
misc_feature
59081..68164
/note="assembly_fragment"
misc_feature
68265..94625
/note="assembly_fragment"
misc_feature
94726..102564
/note="assembly_fragment"
misc_feature
102665..116764
/note="assembly_fragment"
misc_feature
116865..136408
/note="assembly_fragment"
misc_feature
136509..156380
/note="assembly_fragment"
misc_feature
156481..180631
/note="assembly_fragment"

misc_feature 156481..180631
/note="assembly_fragment"
misc_feature 180732..190358
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 45173 a 48605 c 50635 g 44441 t 1504 others
ORIGIN

alignment_scores:
Quality: 549.50 Length: 192
Ratio: 4.504 Gaps: 2
Percent Similarity: 63.542 Percent Identity: 60.417

alignment_block:
US-09-245-198a-2 x AC016876/rev ..

Align seg 1/1 to reverse of: AC016876 from: 1 to: 190358

100 ALaGlyValAspGlyThrValSerGlyTrpGluGlnThrIleAsnSe 116
:::|||||
83375 TCAGGTGTGACGGACGACGAGTGTGCTGGAGGAGGACAGATCAACAG 83326
116 rSerSerProLeuArGlyrAspArGlnIleGlyGluPheThrValIleA 133
|||||
83325 CTCACGCCCTCTGGCGCTACACCGCCAGATCGGGAGTTTATGTCACC 83276
133 rGAlaGlyLeuTrpTrpLeuTrpCys.Gln..... 142
|||||
83275 GGGCTGGGCTCTACTACCTGACTGTCTGACGTAAAGCCCATCTGCTGCAT 83226
142 142
83225 GGGTAACCCACTAAAGACGTGGCAGAGGTTTCCAGAGAGTGGGGAC 83176
142 142
83175 AAGCTACAGGGCTGGAGAGGTGAGTTGGGGTTTGGGTGGATGGATGCC 83126
142 142
83125 TGCCTGCTGAGAAATTGAAATTGAGCGAGGCGAGGAGGAGGCTGG 83076
143ValHisPheAspGlnIleValIleVal 151
83075 ACTGGGCTGTGTTCACCCACCCAGGTGCTGTGATGATGGAGGAGGCTG 83026
151 aLTYrLeuIysLeuAspLeuValAsnGlyValIleValIleValAsnGys 167
|||||
83025 TCTACCTGAACCTGACCTGTGCTGGATGGTGTGCTGTGCTGCGCTGCG 82976
168 LeuGluGlnPheSerAlaThrAlaIleSerSerProGlyProGlnLeuAr 184
|||||
82975 CTGGAGGAATTCACGACCTGCGGCGAGTTCCTCGGGGCCACGACTCG 82926
184 gLeuGlyGlnValSerGlyLeuLeuProLeuArProGlySerSerLeuA 201
|||||
82925 CCTGTGCCAGGTGTGTGTGTGTGGCCCTCGGGCCAGGATCTCTCCCTGC 82876
201 rGLeArGThrLeuProTrpAlaHisLeuIysAlaAlaProPheLeuThr 217
|||||
82875 GGATTCGACACCTCCCTCGGGCCCATCTCAAGGCTGCCCTTCTCTCAC 82826
218 TyrPheGlyLeuPheGlnValHis 225
|||||
82825 TACTTCGACACTCTTCAGAGTTTAC 82802
seq_name: gb_hhg:AC098923
seq_documentation_block: 179030 bp DNA 1linear HTG 20-DEC-2001
LOCUS AC098923
DEFINITION Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

***, 68 unordered pieces.
AC098923
AC098923.4 GI:17973852
HTG: HTGS PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE
AUTHORS

1 (bases 1 to 179030)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbata,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Chen,Z., Chokhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Denny,G., Ding,Y., Ding,H.H., Douthett,K.R., Delgado,O.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Prantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homs,F., Howard,J., Huber,J., Hulik,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Jondah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J.,
Kovar,C., Kravovic,J., Kuresh,A., Landy,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Lousaged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,N., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,
Ogih,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokoken,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitai,N.,
Slason,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Uemari,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Wenstock,G. and Glibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 179030)
Direct Submission
Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064677.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GTOK
Center clone name: CH230-154B15
Assembly program: Phrap; version 0.990329first call to
flndphrap.plst

Consensus quality: 137096 bases at least Q40
Consensus quality: 145080 bases at least Q30
Consensus quality: 152950 bases at least Q20
Estimated insert size: 138980; sum-of-Configs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-Configs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a working draft sequence. It currently
consists of 68 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1	7467	contig of 7467 bp in length
7468	7567	gap of unknown length
7568	12115	contig of 4548 bp in length
12116	12215	gap of unknown length
12216	12313	contig of 8098 bp in length
12313	20314	gap of unknown length
20314	20413	gap of 5176 bp in length
20413	25589	contig of 5176 bp in length
25589	25689	gap of unknown length
25689	31434	contig of 5745 bp in length
31435	31534	gap of unknown length
31535	37595	contig of 6061 bp in length
37596	37695	gap of unknown length
37696	42172	contig of 4477 bp in length
42173	42272	gap of unknown length
42273	46234	contig of 3962 bp in length
46235	46334	gap of unknown length
46335	50605	contig of 4271 bp in length
50705	50705	gap of unknown length
50706	55202	contig of 4497 bp in length
55203	55302	gap of unknown length
55303	58407	contig of 3105 bp in length
58408	58507	gap of unknown length
58508	61286	contig of 2779 bp in length
61287	61386	gap of unknown length
61387	64467	contig of 3081 bp in length
64468	64567	gap of unknown length
64568	68498	contig of 3931 bp in length
68499	68598	gap of unknown length
68599	73449	contig of 4851 bp in length
73450	73450	gap of unknown length
73451	77402	contig of 3853 bp in length
77403	77502	gap of unknown length
77503	81027	contig of 3525 bp in length
81028	81127	gap of unknown length
81128	84144	contig of 3017 bp in length
84145	84244	gap of unknown length
84245	86680	contig of 2436 bp in length
86681	86780	gap of unknown length
86781	89210	contig of 2430 bp in length
89211	89310	gap of unknown length
89311	92339	contig of 3029 bp in length
92340	92439	gap of unknown length
92440	94951	contig of 2512 bp in length
94952	95051	gap of unknown length
95052	98586	contig of 3335 bp in length
98587	98686	gap of unknown length
98687	100597	contig of 1911 bp in length
100598	100697	gap of unknown length
100698	102766	contig of 2069 bp in length
102767	102866	gap of unknown length
102867	105237	contig of 2371 bp in length
105238	105337	gap of unknown length
105338	107305	contig of 1968 bp in length
107306	107405	gap of unknown length
107406	109057	contig of 1652 bp in length
109058	109157	gap of unknown length
109158	111454	contig of 2297 bp in length
111455	111554	gap of unknown length
111555	113704	contig of 2150 bp in length
113705	113804	gap of unknown length


```
* 113805 115933: contig of 2129 bp in length
* 115934 116033: gap of unknown length
* 116034 118193: contig of 2160 bp in length
* 118194 118293: gap of unknown length
* 118294 120414: contig of 2121 bp in length
* 120415 120514: gap of unknown length
* 120515 122311: contig of 1797 bp in length
* 122312 122411: gap of unknown length
* 122412 125629: contig of 3217 bp in length
* 125629 125728: gap of unknown length
* 125729 128345: contig of 2817 bp in length
* 128346 128546: gap of unknown length
* 128547 130563: contig of 1918 bp in length
* 130564 130663: gap of unknown length
* 130664 132945: contig of 2282 bp in length
* 132946 133045: gap of unknown length
* 133046 134356: contig of 1511 bp in length
* 134357 134656: gap of unknown length
* 134657 136478: contig of 1822 bp in length
* 136479 136578: gap of unknown length
* 136579 138159: contig of 1581 bp in length
* 138160 138259: gap of unknown length
* 138260 140456: contig of 2197 bp in length
* 140457 140556: gap of unknown length
* 140557 142544: contig of 1988 bp in length
* 142545 142644: gap of unknown length
* 142645 143784: contig of 1140 bp in length
* 143785 143884: gap of unknown length
* 143885 145304: contig of 1420 bp in length
* 145305 145404: gap of unknown length
* 145405 147391: contig of 1987 bp in length
* 147392 147491: gap of unknown length
* 147492 148908: contig of 1417 bp in length
* 148909 149008: gap of unknown length
* 149009 150147: contig of 1139 bp in length
* 150148 150247: gap of unknown length
* 150248 151752: contig of 1505 bp in length
* 151753 151852: gap of unknown length
* 151853 153441: contig of 1589 bp in length
* 153442 153541: gap of unknown length
* 153542 154953: contig of 1412 bp in length
* 154954 155053: gap of unknown length
* 155054 156244: contig of 1191 bp in length
* 156245 156344: gap of unknown length
* 156345 158095: contig of 1751 bp in length
* 158096 158195: gap of unknown length
* 158196 159309: contig of 1114 bp in length
* 159310 159409: gap of unknown length
* 159410 160682: contig of 1273 bp in length
* 160683 160782: gap of unknown length
* 160783 162602: contig of 1820 bp in length
* 162603 162702: gap of unknown length
```

alignment_scores:

```
Quality: 425.50      Length: 176
Ratio: 4.052         Gaps: 2
Percent Similarity: 59.659   Percent Identity: 55.114
```

alignment_block:

```
US-09-245-198a-2 x AC098923/rev ..
```

```
Align seg 1/1 to reverse of: AC098923 from: 1 to: 179030
```

```
101 GtlyalaspGlyThrValSerGlyTyrGlnLuhThrsIleasnSerSe 117
|||||
26419 GGTGTGATGGACAGTGTAGTGGGAGAGACCAAAATCAACAGCTC 26370
|||||
117 rSerProLeuArgTyrAspArgGlnIleGlyLuphThrValIleaG 134
|||||
26369 CAGCCCTCTGCGTATGACCCCAATTTGGGAAATTCAGTCAACAGG 26320
|||||
134 laglyLeuTyrTyrLeuTyrCys..... 141
|||||
```

```
26319 CTGGGCTCTACTACCTGACTGTCAGTAAGCCCTGGGCTCCATGGGTAG 26270
141 ..... 141
26269 AGCCATGGCTAAGGAGAGAGAGCTTGGCAGAAGATGGTGGAGTGGA 26220
141 ..... 141
26219 GAACCTGGGTTTCATGAAGAAGATGCTGGATTTCCATGAGGCGAGCA 26170
142 .....GlnValHisPheAspGlu 147
26169 GAGCTGTGATTTTGTCTGTCTCTCTGCTCCAGGTCACCTTGATGAG 26120
148 GlyLysAlaValIlyrLeuLysLeuAspLeuValAsnGlyValLeuAl 164
26119 GGGAAAGGACACTACTACGTAAGCTGACTGTGTGTAAGTGTGCTGC 26070
164 aLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerProGlyP 181
26069 CCTGCACCTCCTGGAAAAATATACAGCCACAGTACCGAGATCTCTGGGC 26020
181 roGlnLeuArgLeuCysGlnVal_SerGlyLeuLeuProLeuArgProG 197
26019 CCCAGATCCGTTTGGCCAGCCGCTGCTGCTGTGGATCCGCGCCGG 25970
197 ySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaIaP 214
25969 GTCCTCACTTTCAGACCCGCCACCCCTTAACCTCCCTTAAGCCTCCGC 25920
214 ropheLeuThrTyrPheGlyLeuPhe 222
25919 CCGTCCATTCGCCCGCGGACTCTTT 25894
seq_name: gb_ov:GGA243435
seq_documentation_block:
LOCUS GGA243435 892 bp mRNA linear VRT 28-JUN-2000
DEFINITION Gallus gallus mRNA for putative CD154 (CD40 ligand).
ACCESSION AJ243435
VERSION AJ243435.1 GI:8977828
KEYWORDS CD154 gene.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 892)
Tregaskes,C.A., Young,J.R. and Burnside,J.
cloning of a putative chicken CD40 ligand
unpublished
2 (bases 1 to 892)
Tregaskes,C.A.
Direct Submission
Submitted (24-JUN-1999) Tregaskes C.A., Immunopathology, Institute
for Animal Health, High St, Compton, Nr Newbury, Berkshire, UNITED
KINGDOM
FEATURES
source
Location/Qualifiers
1..892
/organism="Gallus gallus"
/strain="white leghorn Line 0"
/db_xref="taxon:9031"
/clone="chcd4012.2"
/tissue_type="spleen"
73..891
/gene="CD154"
73..891
/gene="CD154"
/codon_start=1
/product="putative CD154 (CD40 ligand)"
/protein_id="CAB95748.1"
/db_xref="GI:8977829"
/translation="MNEAVSPAAPRPMSGSTPSITMKMFQCLSVEMVQVIGTIVLCL
```


repeat_region	/note="MER53 repeat: matches 1. .186 of consensus" 3255. .3292	repeat_region	/note="MER57A repeat: matches 1. .434 of consensus" 23150. .23658
repeat_region	/note="19 copies 2 mer tt 92% conserved" 3878. .3954	repeat_region	/note="TRIGGER1 repeat: matches 774. .1326 of consensus" 23655. .24186
repeat_region	/note="L2 repeat: matches 2164. .2237 of consensus" 4113. .4282	repeat_region	/note="TRIGGER1 repeat: matches 1. .547 of consensus" 24920. .25222
repeat_region	/note="MER5B repeat: matches 4. .174 of consensus" 4285. .4436	repeat_region	/note="AluY repeat: matches 1. .303 of consensus" 25712. .25869
repeat_region	/note="MER5A repeat: matches 1. .174 of consensus" 5142. .5196	repeat_region	/note="MER97c repeat: matches 16. .179 of consensus" 26054. .26081
repeat_region	/note="MIR repeat: matches 201. .255 of consensus" 5206. .5305	repeat_region	/note="14 copies 2 mer tt 100% conserved" 27311. .27518
repeat_region	/note="MER5A repeat: matches 9. .112 of consensus" 5298. .5331	misc_feature	/note="match: STS: Em:G04517" 27350. .27802
repeat_region	/note="MER5A repeat: matches 43. .75 of consensus" 5442. .5744	repeat_region	/note="match: GSS: Em:AQ614744" 27427. .27633
repeat_region	/note="AluSc repeat: matches 1. .305 of consensus" 5965. .6254	repeat_region	/note="MIR repeat: matches 2. .204 of consensus" 27895. .28024
repeat_region	/note="L2 repeat: matches 2423. .2707 of consensus" 6560. .6856	repeat_region	/note="MIR repeat: matches 111. .246 of consensus" 29211. .29252
misc_feature	/note="AluJb repeat: matches 2. .300 of consensus" complement(6576. .7026)	repeat_region	/note="21 copies 2 mer at 83% conserved" 29268. .29531
misc_feature	/note="match: GSS: Em:B63556" 7367. .7776	repeat_region	/note="Cheshire repeat: matches 2216. .2464 of consensus" 29499. .29926
repeat_region	/note="match: GSS: Em:B32322" 7640. .7672	misc_feature	/note="match: GSS: Em:AQ617879" 29533. .29681
repeat_region	/note="MER5A repeat: matches 43. .75 of consensus" 7665. .7771	repeat_region	/note="MUT1D repeat: matches 1. .154 of consensus" 29820. .30127
repeat_region	/note="MER5A repeat: matches 4. .112 of consensus" 8989. .9080	repeat_region	/note="MLTIE repeat: matches 255. .560 of consensus" 30142. .30804
repeat_region	/note="MIR repeat: matches 51. .142 of consensus" 9189. .9701	repeat_region	/note="Cheshire repeat: matches 1528. .2173 of consensus" 30805. .31116
misc_feature	/note="match: GSS: Em:AQ760102" 9190. .9605	repeat_region	/note="AluX repeat: matches 11. .312 of consensus" 31117. .31175
repeat_region	/note="match: GSS: Em:B90415" 9572. .9811	repeat_region	/note="Cheshire repeat: matches 1469. .1528 of consensus" 31187. .32501
misc_feature	/note="MIR repeat: matches 3. .247 of consensus" complement(10614. .11080)	repeat_region	/note="Cheshire repeat: matches 1. .1360 of consensus" 33567. .33766
repeat_region	/note="match: GSS: Em:AQ571033" 12183. .12226	repeat_region	/note="MIR repeat: matches 48. .261 of consensus" 34845. .35149
repeat_region	/note="Alu repeat: matches 253. .298 of consensus" 12235. .12348	repeat_region	/note="AluSq repeat: matches 1. .303 of consensus" 37351. .37408
repeat_region	/note="L2 repeat: matches 2636. .2749 of consensus" 13889. .13938	repeat_region	/note="29 copies 2 mer aa 72% conserved" 37423. .37571
repeat_region	/note="25 copies 2 mer tt 80% conserved" 14485. .15124	repeat_region	/note="MER5A repeat: matches 2. .188 of consensus" 37565. .37656
misc_feature	/note="CpG Island" /evidence=not_experimental	repeat_region	/note="MER5A repeat: matches 1. .105 of consensus" 37723. .38025
repeat_region	/note="59 copies 2 mer cc 62% conserved" 14763. .14880	repeat_region	/note="AluX repeat: matches 8. .308 of consensus" complement(37990. .38488)
repeat_region	/note="59 copies 2 mer cc 62% conserved" 15224. .15285	repeat_region	/note="match: STS: Em:LI18341" 38062. .38113
misc_feature	/note="31 copies 2 mer gg 72% conserved" complement(15328. .15962)	repeat_region	/note="26 copies 2 mer at 92% conserved" /note="complement(38539. .39048)
misc_feature	/note="match: GSS: Em:AQ631047" 15510. .15943	repeat_region	/note="match: GSS: Em:AQ570301" 39057. .39524
repeat_region	/note="match: GSS: Em:AQ126356" 15744. .15771	misc_feature	/note="match: GSS: Em:AQ778077" 39069. .39496
misc_feature	/note="14 copies 2 mer tg 89% conserved" 15972. .16456	repeat_region	/note="match: GSS: Em:AQ441897" 39334. .39389
misc_feature	/note="match: GSS: Em:AQ418183" 15991. .16438	repeat_region	/note="MIR repeat: matches 90. .145 of consensus" 40906. .41237
misc_feature	/note="match: GSS: Em:AQ815493" complement(16128. .16449)	repeat_region	/note="LIME1 repeat: matches 5817. .6122 of consensus" 41485. .41800
repeat_region	/note="match: STS: Em:HS015XG5" 16198. .16233	repeat_region	/note="AluY repeat: matches 1. .310 of consensus" 41802. .41839
repeat_region	/note="21 copies 2 mer gt 97% conserved" 17613. .18726	alignment_scores:	Quality: 109.50 Length: 221
repeat_region	/note="LIMC3 repeat: matches 6631. .7735 of consensus" 18812. .18936	Ratio:	1.117 Gaps: 12
repeat_region	/note="LIMC4 repeat: matches 6487. .6609 of consensus" 21659. .22737	Percent Similarity:	44.344 Percent Identity: 24.887
repeat_region	/note="TRIGER1 repeat: matches 1324. .2412 of consensus" 22738. .23149	alignment_block:	

US-09-245-198a-2 x AL353138/rev ..

Align seg 1/1 to reverse of: AL353138 from: 1 to: 108967

```

42 ArgGluPro...ProGluLeuAsnProGlnThrGluGluSerGlnAsp 57
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14628 CGGAGCCGGACACCGCTCTTTCTTCACCGCTGCCGACGCTCTGAC... 14581
57 lValProPheLeuGluGlnLeuValArgProArgSerAlaPro... 72
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14580 .....ATCACCTTCGGCTCAGCGCGCTGCAACCGCTTC 14547
73 ....LysGlyArgLysAlaArgProArgAlaIleAlaIleHisTyr 87
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14546 CAGCCGAGCGGAGCGCTGCGAGCGCTGCGCTGCGCTC... 14512
88 GluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly... 101
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14511 .....CAGCCACAGCTGCGCAAGATCGAGCGCGCTGCGACGACGCC 14468
102 .....ValAspGlyThrValSerG 108
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14467 TTTTCCCGCTGAGCAGTCACAAAGCCAGCTGTATTCGACTTACACACC 14418
108 LysThrGlu...GluThrLysIleAsnSerSerSerPro... 119
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14417 CCTGGAGGTATGCGAGGCTCTCTCTCCCAACCGGAGCAATAT 14368
120 .....LeuArgTyrAspArgGlnIleGlyGluP 129
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14367 TTGTCCTCCAGGCTGAGCGGATCTCTGCAAGCCAGCGCGCTGGA 14318
129 heThrValIleArgAlaGlyLeuTyrTyrLeuTyr...CysGln 142
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14317 TGAAGGTGAACCAAGGCTATATTAATTAATTTCCATCTGCTGTG 14268
143 ValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuVal 159
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14267 TTGAGAAATCAGAGTAAAGAAATGCGTG...CTGATGATATTTAT 14227
159 lAsnGlyValIleLeuAlaLeuArgCysLeuGluPheSerAlaThrAla 176
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14226 TTCA.....GTGCTAAGAGATCTCAGGACGCTGCCG 14195
176 lAsnSer...ProGlyProGln 182
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14194 TCTCCAAATGAACAACAGCAGGCTGCAAGTGAATGCCAACACTAAA 14145
183 LeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGly... 197
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14144 TTAAAGTATGCTCTTCTGCAATCTTGCATTAAGAGAGCGACACT 14095
198 .....SerSerLeu 201
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14094 TCATTTCTTGCCCTAAACAAGCAAAAGAAATGCAAGGCTCATCTTAA 14045
201 rglLeuArgThr 204
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14044 GACTCAGAAAC 14034

```

seq_name: gb_ba:SC1C2

seq_documentation_block:

LOCUS SC1C2 42210 bp DNA linear BCT 15-JAN-1999
 DEFINITION Streptomyces coelicolor cosmid 1C2.
 ACCESSION AL031124
 VERSION AL031124.1 GI:3355667
 KEYWORDS 3-isopropylmalate dehydratase large subunit; 3-isopropylmalate dehydratase small subunit; 3-isopropylmalate dehydrogenase; branched-chain amino acid aminotransferase; carboxyl transferase; delta-1-pyrroline-5-carboxylate dehydrogenase; glx; glutamyl-tRNA synthetase; histone-like DNA binding protein; hydrolase; lve; leuB; leuc; leuD; lyase; secreted lyase; transfer-RNA-Gln;

SOURCE

transfer-RNA-Glu: ureaB; urease alpha subunit; urease beta and gamma subunits; ureC.
 Streptomyces coelicolor A3(2).
 Streptomyces coelicolor A3(2)

ORGANISM

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE

1 (bases 1 to 42210)
 Murphy, L. and Harris, D.
 Unpublished

REFERENCE

2 (bases 1 to 42210)
 Parkhill, J., Barrell, B.G. and Rajandream, M.A.

TITLE

Direct Submission
 Submitted (23-JUL-1998) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

REFERENCE

3 (bases 1 to 42210)
 Redenbach, M., Kleier, H.M., Denapalte, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.

AUTHORS

A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome

JOURNAL

Mol. Microbiol. 21 (1), 77-96 (1996)

MEDLINE

97000351

COMMENT

Notes:
 Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).

FEATURES

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 1C2 lies between BD9 and 7A1 on the AseI-B genomic restriction fragment.

Location/Qualifiers

1..42210
 /organism="Streptomyces coelicolor A3(2)"
 /strain="A3(2)"
 /db_xref="taxon:100226"
 /clone="cosmid 1C2"

gene

1..1228
 /gene="SC1C2.01"

CDS

/gene="SC1C2.01"
 /note="SC1C2.01, probable delta-1-pyrroline-5-carboxylate dehydrogenase, partial CDS, len >408 aa; similar to many e.g. TR:O50443 (EMBL:AL010186)
 delta-1-pyrroline-5-carboxylate dehydrogenase (M. tuberculosis) (543 aa), fasta scores: opt: 1753 z-score: 2214.2 E(0): 0, 65.1% identity in 407 aa overlap, and PUT2_HUMAN delta-1-pyrroline-5-carboxylate dehydrogenase

```

1. 563 aa), fasta scores; opt: 1273 z-score: 1730.4 E(): 0,
2. 48.5% identity in 408 aa overlap. Contains P500687 and
3. P500070 Aldheyde dehydrogenases glutamic acid active site
4. and cysteine active site and Pfam match to entry PF00171
5. aldehyd. Aldheyde dehydrogenases, score 114.70, E-value
6. 1.8e-30"
/codon_start=2
/transl_table=11
/product="delta-1-pyrroline-5-carboxylate dehydrogenase"
/protein_id="CA11968.1"
/db_xref="GI:3355668"
/db_xref="SPTRMBL:086502"
/translaltion="IDSPCEIDPFMRPNVYHARNILAEQPPANSGVNNRMDHRPLEG
FVYATTPNFSALANLPTPALMGVYVWVWKSPTQTHAVALMLQLEAGLPGVINGV
LVGTGJALFVALEHNDLHGTSTPTFOIMAKTVNNIEKTRTPYLVEGEGG
DELVAHSGADAVIKVITFGAEGEGGCGKCSSTRAYIPASITMGNGFEFEFAEVDYIT
TMGVDTDSNFTGAYIDERSFANKKAIDBRAKEDCTTIVAGGSIDSGEGFVRPPTVY
ECTPENSLETFEFGFFLAHVHYDDSDADADAMLTQWESYSDALGVSYSNDRAAP
AAYTMETKIRYVAAAGNFIINDKSTGAVVGOQPFGGGRASGTNDKAGAPOMIMETLTLRAL
KETVLAADPDAVYPPHMG"
1. .3807
/ncote="true overlap with cosmid 8D9"
2. .1189
/ncote="SC1C2.01"
/ncote="Pfam match to entry PF00171 aldehyd, Aldheyde
dehydrogenases, score 114.70, E-value 1.8e-30"
464. .487
/ncote="SC1C2.01"
/ncote="P500687 Aldheyde dehydrogenases glutamic acid
active site"
548. .583
/ncote="SC1C2.01"
/ncote="P500070 Aldheyde dehydrogenases cysteine active
site"
1246. .1283
/ncote="hairpin_loop with 18bp stem"
1522. .1527
/ncote="Possible RBS upstream of SC1C2.02"
1537. .1668
/ncote="SC1C2.02"
1537. .1668
/ncote="SC1C2.02"
/ncote="SC1C2.02"
/ncote="SC1C2.02, questionable ORF, len: 43 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1C2.02"
/protein_id="CA11969.1"
/db_xref="GI:3355669"
/db_xref="SPTRMBL:086503"
/translaltion="MAETTVRRVRHLSRFTSDSRKNMAAALQRLSDLRDNGCATGH
1744. .2787
/ncote="leuB"
/ncote="leuB"
1744. .2787
/ncote="SC1C2.03, leuB, probable 3-isopropylmalate
dehydrogenase, len: 347 aa; similar to e.g. LEU3 CORGL
3-isopropylmalate dehydrogenase (EC 1.1.1.85) (340 aa),
fasta scores; opt: 1450 z-score: 1318.1 E(): 0, 68.0%
identity in 338 aa overlap. Contains P50470 Isocitrate
and isopropylmalate dehydrogenases signature and Pfam
match to entry PF00180 isodh, Isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
/codon_start=1
/transl_table=11
/product="3-isopropylmalate dehydrogenase"
/protein_id="CA11970.1"
/db_xref="GI:3355670"
/db_xref="SPTRMBL:086504"
/translaltion="MKSMLAVATIPEDGIGQEVVAEGKLVLSAVLPDYVKLETKEFDE
GARRHATGELTDADLADKADADLILDAIGDPSPSSVERGLFLKLRAPFDHVVV
IRPSKLTLPGVAITPLAGQPELDIVVVEGTEGATIGPVGNGGTRGHEVAETEVSNTAA
EVEVRVDAEFRAQDARRKTLTLVHNKNVLTLEGGHILMTIIFPKVAAEPEVYTDYLIHNI

```

```

misc_feature
|note="Leub"
|gene="leub"
|note="Plan match to entry PF00180 isodh, Isocitrate and isopropylmalate dehydrogenases, score 356.60, E-value 1.3e-105"
2458..2517
/gene="leub"
/misc_feature
|note="PS00470 Isocitrate and isopropylmalate dehydrogenases signature"
3038..3042
/note="possible RBS upstream of lve"
3052..4140
/gene="lve"
3052..4140
/gene="lve"
/note="SCIC2_04, lve, probable branched-chain amino acid aminotransferase, len: 362 aa; similar to many, e.g. lve_BACSU putative branched-chain amino acid aminotransferase (362 aa), fasta scores: opt: 1047 z-score: 1401.5 E(): 0, 44.7% identity in 360 aa overlap. Contains PS00770 Aminotransferases class-IV signature"
/codon_start=1
/trans_table=1
/product="branched-chain amino acid aminotransferase"
/protein_id="CAI19971.1"
/db_xref="GI:3353671"
/dx_xref="SPRMBL:O86505"
/translation="MTPTTELKPSAHPLSDSERAAILANPGFGRHFTDHWIKTKTEGRNSIDOLVYARSLDPATPMVLVAHQEIEGKAVRPDSVATRRPKNARGAFOA SRRRLGMELVDVFETFCALVADKKQVPAHGSESYLRPMITAEVGIGVRPAN EFLDTIASPRAGATFRPGSVRSVEDRRAVPGGMGDAKTGTNAAISLAQEA AKGDQVCYLDAIERKWVEELGKMLFVGNKTVLSLTESIIEGVTRDSLTVARD LGYAEGSRVSDVMORDSENGTLFEVACSTAAVITPVGVKRAQAMOOOSGETGE VTORLRNALDLIORGTADVDRPHGMTLA"
3763..3852
/gene="lve"
/note="PS00770 Aminotransferases class-IV signature"
/complement(4214..5644)
/gene="SCIC2_05c"
/complement(4214..5644)
/gene="SCIC2_05c"
misc_feature
|note="PS00770 Aminotransferases class-IV signature"
/complement(4214..5644)
/gene="SCIC2_05c"
/complement(4214..5644)
/gene="SCIC2_05c"
alignment_scores:
Quality: 108.50 Length: 225
Ratio: 0.882 Gaps: 13
Percent Similarity: 54.667 Percent Identity: 29.333
alignment_block:
US-09-245-198A-2 x SCIC2/rev ..
Align seg 1/1 to reverse of: SCIC2 from: 1 to: 42210
19 SerLeuGlySerTyrAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
:::||||||::: |||:::||::|::: |::: |:::|:::|:::
9451 GGCTCGGGCGGCGAGGCCCTCCGCTCCACTCCAGGTGCGAGCAGA 9402
35 u.....LeuThrAlaGluAspArgArgGluPro.ProGluLeuAsnPro 49
::: ||| ||| ||| ::||| ||| ||| ||| |||
9401 GTTCGCCTTGAGCGCGAAGTAATAACAAGAGTGGCGGCTGCTCAATCCG 9352
50 GluThrGlnGluSerGlnAspValValProPheLeuGluGlnLeu.ValA 66
::::: ||| ::||| ||| :||| |||
9351 ACCTCGGCGCGGCGCGCCCATGTGACCTTC.....TCCAGCCC 9311
66 rPrProArgArgSerAlaProLySGlyArgLyAlaArgProArgArGa 82
|| ||| ||| ||| ::||| ||| ||| ||| ||| |||
9310 GGCGTGGCGATCATCTCCATGGCGGCGAGCACGCTCGCGCGCG 9261
83 .....IlalaAlaHisTyrGluValHisProArgProGlyGlnaspG 97
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
/protein_id="AAG05154.1"
/translacion="MRASIVCYFSLALPASSPGMASGGRYLTFWVNDNGVHNTFVDS
QFDRROARORIGOSDORALINKADATQMPGNSGSGESKRRTYTTIDANGVNSF
YAGAQAAGRADYVLPNGEHAIEYIDAEATEGKFVSENGSPRTYTWDEKGRHNSQ
IPDEBARAFKREAPSSQIAFTEGROIIEFKORPALPLGDGAQOSDALKLKTE
MTIDLDYQDQRCCQIASEDFEILDADPRYLELRNFSFSDPFPMGRSFYAIRLP
RSQGVYGLRIRSFANQVYPSLFLDEHKRPRLVSDAVYKLNPTWRYAFTEGTV
PVARNRGERVYLLLTDEDSLQTLDNKPKRPLQDLAVNBAQMVHEHAROGGFELA
VVR"
complement(5366..6349)
/gene="PA1766"
complement(5366..6349)
/gene="PA1766"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG05155.1"
/db_xref="GI:9947745"
/translacion="MSIMKTMKDLTAKVANGINRRNADYVLYKRNRLHPVVDKITT
KERALEAGIHVEYGGIIEETEKIDKNEIIEGHNDPVIKPAQAGDGIWVADPRE
GRKTVSGKIVSHHEIEPHHISLISGLXSLGHRDRLIEYRTPDOIRKISYEGP
DIRIYIMGTPYAMALRLPTROSGRANLHOGAIGCVVDLATCYTTLGTVLNNKIAH
PDTNADVQVQLPMDGFMKLAAGCYELSGLYIGDMVLDQEKGPLLELNARPLN
IQIANDSGLVHRCHAVEARLEELKAKGIEBSAEBERYFSOOLFGHVOOHOG"
complement(6353..7879)
/gene="PA1767"
complement(6353..7879)
/gene="PA1767"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG05156.1"
/db_xref="GI:9947746"
/translacion="MRSLNHLKYLIALILTLGVAITAYOIFVOCIVYTEARDDDLNN
IDAKVDVATPKTPVKQVMPPLSDYSLNSFVSNNGVSNIRADGNRKVWSAR
RANGKQFLYRLVLTFRYTGEEKPIFRDSPLLEGPERIAEALLAP1ROHSADY
ETVYSEAIKIRANNNDNDNVLKMGDTSLERKKAHIVETLSIAHPMERHTILVAD
LOOSPMLRSPNGENMLFYNVYTGEOGLPSDRILIMLGDPELMTIDGGKAOVSFL
NSSEMAIRLAKLSDENTAEAFLEYSXGLPLISOQFOQIMIMPIEIVYLILRNLI
GLDITLFTFVYLALARETGLOMGLFVITRILGSLSYLHLLQMLPRISVVL
TFVYVMIAVLSLFSHKLGFESGLSIALFPVILMTLTERLSITMEERGSGHAMVAIG
TLFPAASLAHILMRIPELVYFIETFPVALLIVAFMLMGRYRGYRLTELFRKALLD
"
complement(7885..8424)
/gene="PA1768"
complement(7885..8424)
/gene="PA1768"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG05157.1"
/db_xref="GI:9947747"
/translacion="MKLKLPLALPLCLFIAVPGLSVAEKTVYGLNEXARINDPDIQA
AKLDTGAKTASLSLRDIKRRRGETWRYLATDNADDP1IEKPLARISKIRHSD
FNDEGKATYARAPIELQVCMGAIRTIENVLDRSAFYQPLIGSALKKFDALVDP
SLKYSACKPGCKDPKAPAE"
complement(8715..9539)
/gene="PA1769"
complement(8715..9539)
/gene="PA1769"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG05158.1"
/db_xref="GI:9947748"
/translacion="MQMKRTAFISDGTGTAETLIGSLAOFENISFVXLTREYIDT
EERAKAVQOINNAESDGAIP1FDIIVNDIAYVAQSGFMID1FAFPLSLEDE
LSADSSYVSGKSHGNSNMDBRIEAVNLDDKQDARHHYKALDILIVGSRCK
TPRCIYALQYGRIPAAVYPLTEEDMERLQPLNALKQKHKHLLFGTIDPDRLTRNR
KPNRSYASFQCEFEVREVSFLRRENIATINSTHSEVEISAKTILVEKVEREK"
9705..12080
/gene="psa"
```

```
CDS
/note="PA1770"
9705..12080
/gene="psa"
/codon_start=1
/transl_table=11
/product="phosphoenolpyruvate synthase"
/protein_id="AAG05159.1"
/db_xref="GI:9947749"
/translacion="MVEYVVSLLDKLVHDVDEHVGKRNASIGEMISNLGAGVSYPGE
ATTAQARDFLEDSGLNDRTHAALDADLDVYNALATGQIQRWYAEAPRLDSE
IRQAFALANGNDNLAVRNSATAEPLDPAESGQDETLNTRGVNVAIRAEVFA
SLENDRAIAYRVHGFDPKLVALSAGQVRWRSFTAGVWFTLDTSGRDVFITG
AYGIGETVVOGAVPDEFYVHKPTLEAGRPALIRNLNYSKAIKM1YDEKAKASVY
VDVDRADRARFELSDAEVTELAQAM1IEKHGSRPMDIEMAKDDGDKLYIVARPET
VKSASATVMERLYLLEKGTIVYNGRA1IGORIGPVYVINDSEMDKVOGDIYSD
MTDPDEPVYKRRASAYITNNGGRTCHAAITARELGPAYVCGCATIIDLQGCYTVS
CAEDGTGFIIEGELGFDVRKNSVDAMPDLPEKIMNVGNPDRAEDFAQLNEGVGLAR
LEFIINRMIGVHPKALNFAGLPADIESVEKRIAGYPDVGFEKVLVSGI1STLAA
FWPKYIVLR1SDFESNEYANLIGKLTYEPEENPMLGFRGASRYISFSDCELECR
ALKVRNEMGLTNYEIMVPPVRLTGEASOVVELLAGNLRGENGLKVIIMMCELPNA
LADPELEFPDGSFGSNDLTOLITGLDRDGIYAHIFDRNRPVYKLLANATACAK
AGKITGICGGPSDHPDLARKIMDQIESVSLNPDVLDTMFLAEGODA"
```

```
alignment_scores:
  Quality: 107.50      Length: 161
  Ratio: 1.361        Gaps: 6
  Percent Similarity: 49.068  Percent Identity: 27.329
```

alignment_block:

US-09-245-198a-2 x AE004602/rev ..

Align seg 1/1 to reverse of: AE004602 from: 1 to: 13243

```
5 GlyLeuAlaLeuAlaCysLeuGluLeuLeuValValSerLeuCl 21
12726 GCGCGACGACGAGCGCGGAGGATGCCG..... 12697
21 yserTPaIatnrlLeuSerAlaGln..... 29
12696 .CCATGGCAGACCTGACACGCCAGTCCGGCTGACTGCTGCCACCGAA 12648
30 ..GluProSerGlnGluLeuThrAlaGluAsp.....ArgArgGlu 43
12647 CAACGCCACGAGAACGATCCGTGCTTCAGATGAGCCCGCCGACGCC 12598
44 ProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPh 60
12597 CCGCCGCCACGACACACCCACCCAGGAGTGGCAATCCAGTGCGCCGCC 12548
60 eLeuGluGlnLeuValArgProArgArgSerAlaPolysGlyArgLyAs 77
12547 TGCCCGCACTGCTCGTGATGAAGAGGCCAGCGCCGACGATCGTAGCG 12498
77 IaArgProArgArgAla..... 82
12497 CCGGTGATGGCGAGACCCGGTGTCCCGTAACCATGCTTGCAATGACA 12448
83 ..... 93
12447 GCCCATGGCCGGCGCATTTCCGCGACGACACATCGAACCCCGCCGCCG 12398
93 oGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSer...GlyT 109
12397 AGTAAAGGACCTAATCCGAGCCGACGCGGAGTACCAAGCGCGGGGTT 12348
109 rPGLuGluThrLysIleAsnSerSerSerProLeuArgTyrAspArgGln 125
12347 GGAATAAATCCCGTGCAGAGATCACAGCGGTCGCCGCGTGCGCCG 12298
126 lIeGlyGluPheThrValIleArgAlaGlyLeu 136
12297 .....CGCCCTGCTGCGAGCGCGGTTA 12274
```

THIS PAGE BLANK (USPTO)

FEATURES	SOURCE	CDS	BASE COUNT	ORIGIN
1..1168	location/Qualifiers		242 a	
/organism="Mus musculus"			360 c	
/db_xref="taxon:10090"			298 g	
/cell_type="peritoneal macrophages"			268 t	
<1..679				
/note="ligand in the TNF family; secreted protein"				
/codon_start=2				
/product="TWEAK"				
/protein_id="AAC53517.1"				
/db_xref="gi:2707221"				
/translation="VLSIGLALACLGLLVVSHGSMATLSAOEPSOEELTAEPREPE PELMPQEEESODVYPRLEFOLYRPRRSPKGRAPRAIAAHYEVHRRPGDQAQACV DGITYSCHMEETKINSSPELRYDROJGEPTVYIAGIYLYLCOVHFEDEGRNVIYKEDLVN GVLLARCLDEFSATFAASPGQRLKCOVSGLLPVRPGSSLRITLPMHLKAAPFLYV GFLQVH"				

Query Match	100.0%	Score 1168;	DB 10;	Length 1168;
Best Local Similarity	100.0%	Pred. No. 1.1e-295;		
Matches 1168; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

Db	721	CAGGAGCATACACACACTCCCTACCCACCCTCCCTCCGCTGCTCTT	780
Qy	781	ggtccagtcctgtctctctctcaaaagcagcagagctgttccatgtttccatccaca	840
Db	781	GGTCAGCTCGTCTCTCTCTCAAAAGGACGAGGCTGTGTTCAGATGTTTCATTGCCACA	840
Qy	841	gaagatcctctgtcctctcttaacatccatccatccacaacaactctcacctcctagctcc	900
Db	841	GACGATCTCTGCTCTTCTTACATCTCCATCCACACAACTTCCACCTCTCATGCTTC	900
Qy	901	caaaagccctactatctccctgacctcccccccacacccaacccgacccagctgttatgact	960
Db	901	CAAAAGCCCTACTTATCTTATCCCTGACTCCCCACCCACTCACCGACACAGTGTATTGACT	960
Qy	961	ctgtgacacagcactagatctgagctgagactgtgtgcaagaagcagaagaacctgagac	1020
Db	961	TTGTGTCAACAGGCACTGAGATGGGCTGACCTGTGTGCAAGAACCAAGAACCTGTGGAC	1020
Qy	1021	taggccagaagcttcccaactgtgagggggaagaagctgtgaggacaagctctccctgatalcc	1080
Db	1021	TAGGCGAAGATTCCTCCAAAGTGTAGGGGGAAGAAGCTGGGGAACAAGTCCCTCCCTGATCC	1080
Qy	1081	ctgtgagatcttggaaagatactatcttattatattatgtgtgacaaaatgtttaaigtatat	1140
Db	1081	CTGTGATTTTGGAAAGATCTATTTTATTATTATTGTGCACAAATGTTAAATGATAT	1140
Qy	1141	taaaagagataatacatgatcttctcttc 1168	
Db	1141	TAAAGAGATAAATCATGATTTTCTCTTC 1168	

RESULT	2				
LOCUS	AC069459/c				
DEFINITION	AC069459	203083 bp	DNA	linear	HTG_27-JUN-2001
SEQUENCE	Mus musculus chromosome 11 clone RP23-168P5,				WORKING DRAFT
ACCESSION	AC069459				
VERSION	AC069459.23	GI:14547768			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE
AUTHORS
1 (bases 1 to 203083)
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
Deerlich, D., Thomas, S., Okwuonu, G., Carlick, C., Garner, T.,
Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,
Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,
Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Hochs, S.,
Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,
Gorell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M.,
Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,
Kovar, C., Liu, J., Liu, W., Louisse, H., Lozado, R. J., Matlin, R.,
Massey, E., McLeod, M. P., Mei, G., Moore, S., Morgan, M., Morris, S.,
Neel, D., Nelson, A., Nguyen, N., Nguyen, N., Ogan, M., Parish, B.,
Perez, L., Relter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,
Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,
Munzy, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,
Worley, K. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 203083)
Worley, K.C.
Direct Submission
Submitted (31-MAY-2000) Human Genome Sequencing Center, Department

COMMENT

On Jun 25, 2001 this sequence version replaced gi:12621264.
Center: Baylor College of Medicine
Genome Center
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>

VECTOR: pBACE3.6
This sequence is the entire insert of clone RP23-422L16.
Location/Qualifiers
1. .234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_11b="RPC1-23"
/complement(84050..84131)
/note="Sequence from uni-directional primer reads and dcrp
big dye terminator reads only."

misc_feature
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

Query Match 60.9%; Score 711; DB 10; Length 234182;
Best Local Similarity 99.6%; Pred. No. 2.1e-175;
Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

```
425 caggctgacatttgatgaggaagagctgtctactgaagctggaactgtgtgtaacggt 484
|||||
66233 CAGGTGCATTTGATGAGGAGAAAGGCTGTCTACCTGAAGCTGGACTGTGTAACGCT 66264
|||||
485 gtgctggccctgcgcctgcctggaagaaattctcaagccacaagcagaagctctccctggccc 544
|||||
66263 GTGCTGGCCCTGCCTGCTGGAAGAAATTCACAGCACAGCAGAAAGCTCTCTGGGGCCC 66204
|||||
545 cagctccgtttgtgcagaggtgtctggcgtgtgtccgcctgcgcgcagaggtctcccttgg 604
|||||
66203 CAGCTCCGTTTGTCAGAGTGTGGGCTGTGGCTGTGGCTGGCCAGGGCTCTCCCTTCGG 66144
|||||
605 aaccgacaccccccctggcctgcctcaacttaagcctgcgccctcccaactcttggagctc 664
|||||
66143 ATCGGCAACCTCCCTGGGCTCATCTTAAGGCTGCCCTCTCTAACCCTATTGGAGATC 66084
|||||
665 ctccaagctcaatgaggggctgtctctccagatctccaaattctccctgtgctcgaag 724
|||||
66083 TTTCAAGTTCACTAGAGGGGCTTGTCTCCAGATTCCTTAACCTTCCCTGGCTCCAGG 66024
|||||
725 agcatcaacacacccctccacaccccccacccctccacccctcgtgtgtctcttggc 784
|||||
66023 AGCATCACACACCTCCCTACCCACCCCACTCTCCACCCCTCGTGTCTCTGGTGC 65964
|||||
785 cagctcctgt-ctctcctcaaaagcagccagagctgtgtcacatg-cttccattccacaga 842
|||||
65963 CAGTCTGTCTCTCTCTCAAAAGCGACGAGCTGTTCACATGTTTTCATTCACAGCA 65904
|||||
843 cgtatcctgtctctc-ctaaatcccaatcccaacacaaactatcccaactcaatgctcc 901
|||||
65903 CGTATCTTGTCTCTCTTAAATCCATCCATCCACACAACTATCCACTCTCACTAGCTCC 65844
|||||
902 aaaaacccctacttaccctgaactcccccacacacacacacacacacacacacacacacac 961
|||||
65843 AAAAGCCCTTACTTATCTCTGACTCCCAACCCACTCACCCGACACACGCTTTATTTAGACTT 65784
|||||
962 tgttcacacagcagctgagatgagctgtgacgtgtgtgcagagacccaagaaactcggagct 1021
|||||
65783 TGTGCACACAGCAGCATGAGATGGGCTGGACCTGGTGGCAGAGACGAGAACTGGGACT 65724
|||||
1022 aggcacaagaattcccaactgtgagggggaagagctgggagacaagctccctccctgatacc 1081
|||||
65723 AGGCACAGAAATTCCCAACTGTGAGGGGGAAGAGCTGGGGAACAAGCTCTCTCCCTGGATCC 65664
|||||
1082 tgttgatcttgaagaagatactatttatttatttgaacaaatgttaattgatatct 1141
|||||
65663 TGTGATTTTGAAGATACATATTTTATTTATTTATTTGACAAATGTTAAATGATATT 65604
|||||
1142 aaagagaataaatcatgattctctctc 1168
|||||
65603 AAAGAGATAATATCATGATTTCTCTCTC 65577
```

RESULT 4
AX201324 1353 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0154486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1353)
Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
Hillan,K.J., Masters,S.A., Pan,J., Piltl,R.M., Roy,M.A., Smith,V.,
Stone,D.M., Watanabe,C.K. and Wood,W.I.
Compositions and methods for the treatment of tumour
TITLE Patent: WO 0153486-A 3 26-JUL-2001;
JOURNAL Genentech, Inc. (US)
FEATURES
source 1. .1353
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN

Query Match 53.8%; Score 628.6; DB 6; Length 1353;
Best Local Similarity 76.2%; Pred. No. 4.7e-154;
Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

```
2 gtgtgagctggcctggcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctg 61
|||||
130 GCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 189
|||||
62 gggagctgggaaagctgtgtctcccaagagcctctcaagagagagctgacagcagaagac 121
|||||
190 GGGAGCGCGGATCGCTGTCCGCGCAGAGGCTCCAGAGGAGCTGTGGCAGAGAG 249
|||||
122 cgcggagagcccccctgaactgaatccccaagacaagaaacccagagatgtgtacatttc 181
|||||
250 GACCAAGACCCGCTGGAACTGAATCCCAAGACAAAGAACCCAGATCCCTGGCCCTTC 309
|||||
182 ttgaaacaactagtcgcgcctcgaagaagtgtctcctaagaagcggaaagcggccttcgc 241
|||||
310 CTGAACCACTAGTTCGGCTCCGCAAGTGCACCTAAAGCCGGAACACGGGCTCGA 369
|||||
242 cgaactcttgagacccattatgagttcatctctctgcgcagagacagagatgagacaaaga 301
|||||
370 AGAGGATCGGACCCATTATGAAAGTTCATCAGAGCTTGACAGAGGAGCGCAGGCA 429
|||||
302 ggtgtgagatgagagatgagctgtgaggaagaagacaaatacaacacagctccagccctctg 361
|||||
430 GGTGTGAGACGGGACAGTGAAGTGGCTGGGAGGAGAACCAAGATCAACAGCTCCAGCCCTCG 489
|||||
362 cgtacgacgcagatgtgggaattacagtcacaggtcaggtgcctactactcttac 421
|||||
490 CGCTAACACCCCGAGATCGGAGATTATAGTTCACCCGGCTGGGCTCTACTACTCTGAC 549
|||||
422 tgttaagtgacatttgatgaggaagaagctgtctcaacctgaagcagagactgtgtgtaac 481
|||||
550 TGTGAGGTGCACTTTGATGAGAGGAGGCTGTCTACCTGAAGGTGAGACTGTGCTGTGAT 609
|||||
482 ggtgtgagccctgcgcctgtccttgaaagaattctccagccacaagcagaagctccctcgg 541
|||||
610 GGTGTGCTGGGCTGTGCTGTGCTGTGAGGAATTTCTACGCACTGTGGGCGAAGTTCCTCGGG 669
|||||
542 cccagagctcgtgtgtgcagagtgctgtgagctgtgtgcgcgtgcgcgcagaggtctccctc 601
|||||
670 CCCAGACTCCGCTCTGCAAGTGTGTGGCTGTGGCCCTGGCCAGGGCTCTCCCTG 729
|||||
602 cggatccgacaccccccctggcctcattcttaagctgcgcctccctcaactacttggga 661
|||||
730 CGGATCCGACACCTTCCCTGGGCGCAATCTCAAGGCTGTGCCCTTCTCACTACTTCCGA 789
```


Db	830	-----CCTGACAGACTCTCTGTGGGACCCGGTCCCTCTTGCCCCACCTCAGCCGCTCTTT	904
Qy	781	ggtccagctcgtcgtcctcc--tcaaaggcagccagagctgtgtccatgttccatcc-	837
Db	905	gctccagacgtcccccctccctctgacagagctgctggcctgttgcacgcttttccatcc	964
Qy	838	-----acagacgttacctgtcctctcttctaacaatccatcccaacaactatccactc	891
Db	965	ACATTAATACATATATTTCCACACTCTTATCTTACAACTCCCCACCCGCACTCTCACCTC	1024
Qy	892	actagctcccaagaagccctactat-----ccctgacctccccaccact	936
Db	1025	ACTTACGTCCTCCCATTCCTCTGACCCCTTTGAGGCCCCGACGTGATCTGCATCTCCCGCCCA	1084
Qy	937	caaccgacacagctgttalttgaactctgtgac-----	968
Db	1085	CAGACCCCGCAGTCTATTTGTGTCTCACTGTACTCTGTGGGCAAGATGGTCCAGAAAGACC	1144
Qy	969	-----caggcactgagatgtgagctgtgacctgtgtgcagaaagccagagaacctgagtag	1023
Db	1145	CACCTCAGGACACTAAGAGGGGCTGGAAGCTGGCGGACAGAACCCAAAGAGACTGGGCGTAG	1204
Qy	1024	gccaaagttcccaactgttgaggggaagagctgtggagcaagctctctcccgga-----tc	1079
Db	1205	GCCAGAGAGTCCCAATGTGAGGGGCGAGA-AACAAGCAAGCTCTCCCTTGAGAAATTC	1263
Qy	1080	ccctgagatcttgaa--agatacattcttaatacttgagacaatgt--taat	1134
Db	1264	CCTGTGAGATTTTAAACAGATATATTTTATTTATTTGTGTGACAAATGTGTATAT	1323
Qy	1135	ggatcttaaagaagaaataatcatga	1159
Db	1324	GGATATTAAATGAAATGAGTCATAA	1348

RESULT 6

AF030099

LOCUS	AF030099	1306 bp	mRNA	linear	PRI 20-DEC-1997
DEFINITION	Homo sapiens TWEAK mRNA, complete cds.				
ACCESSION	AF030099				
VERSION	AF030099.1	GI:2707218			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1306)				
TITLE	Chieneporliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L.				
JOURNAL	TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis				
MEDLINE	J. Biol. Chem. 272 (51), 32401-32410 (1997)				
AUTHORS	2 (bases 1 to 1306)				
TITLE	Bourdon,P., Hession,C., Tizard,R. and Browning,J.				
JOURNAL	Direct Submision				
FEATURES	Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA				
source	Location/Qualifiers				
	1..1306				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="17"				
	/map="17p13"				
	/tissue_type="tonsil"				
	18..767				
	/note="Ligand in the TNF family; secreted protein; start codon not verified experimentally"				
	/codon_start=1				
	/product="TWEAK"				
	/protein_id="AAC51923.1"				
	/db_xref="GI:2707219"				
	/translation="MAARKRSQRGRGRGEPCTALVPLAUGIALGIALACLGILLAVSL"				

[illegible]

Db	985	ACTACTCCCTCCCAATGCCGACACCCTTTAGAGGGCCCCAGTGAATCTGCACCTCCCCTCGGCCA	1044
Oy	937	cacccagcacgctgttatltgaacttltgcac-----	968
Db	1045	CAGACCCCCAGGGCATTGTGTTCACCTGTACTCTGTGGCAAAGATGGGTCCAGAAGACCC	1104
Oy	969	-----caggacatgatatgggcttgaccttgtgtgaggaaagccagaagaccctggatactag	1023
Db	1105	CACCTTCAGGCACTAAAGAGAGGGCTGGACCTGGCGGCGAGAAAGCCAAGACTGGGCTTAG	1164
Oy	1024	gccagaagtctccaactctgtagggaggaaagcgttggagacaagcttcctccctgta----tc	1079
Db	1165	GCCAGGACTTCCCAATGTGAAGGGCGCAGA-AACAAGAACAGCTCTCTCCTTGGAATAATTC	1223
Oy	1080	cctgttgatllttgaaa--agatactattltaattatatlctlgacaaaaatgt--taaat	1134
Db	1224	CCTGTGGATTTTTTAAMACAGATATATTTTTTATTATTGTGACAAAATGTTGATAAAT	1283
Oy	1135	ggaataaagaagaataaatics	1156
Db	1284	GGATTTAATATGATTAAGTCA	1305

[illegible]

OY	362	gcctacgcccgcagaatttgggaattttaaagcatcaggagcgtgagctctactactgtac	421
Db	433	cgcctacAACGCCCAATCGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTAC	492
OY	422	tgctcagtgatcatttgatgaggaaggaagcgtgtactaacttgaagctgtgacttctgtgtgaac	481
Db	453	tgTCAGGTGCACTTTGATGATGAGGGGGAAGGCTGTCTACCTGAAGCTGGACTTGGTGGAT	552
OY	482	ggtgtgtgtgcccctgtgcgtctgtcctgtgaagaattctcaagccaagcaagctctctcctgag	541
Db	553	GGTGTGCTGGCCCTCGCGCTGTGAGGAAATTCACGACCACTGCGGCGAGTTCCCTCGGG	612
OY	542	cccgcagctccgttctgttgcgaagtgatctgaggtgtgttcgagctctcggccaaaggtcttcctt	601
Db	613	CCCCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTTTGCCCTCGGCGCAAGGCTCTCTCCTG	672
OY	602	cgaatccgaacccctccctcctgaggtctcatctlaaagctgtcccccttccactaacttltgga	661
Db	673	CGGATCCGACCCCTCCCGGGCCCATCTCAAGGGCTGCCCTTCCTCCTCACTACTTGGGA	732
OY	662	ctcttccaagttcatctgagggagccttgcctctccagatctcctttaaacttccctgtgcctc	721
Db	733	CTCTTCACAGTTTCATGTAGGGGCGCTGTGCTCCCGCACTCTCCAGGCTTCGCGCTCC	792
OY	722	aggagcatcaccaacacccctccctaccaccaaccacccactctctccacccccctc- -ggtgtgtcct	780
Db	793	-----CCTGTGACAGCTCTCTGTGGGACCCGAGTCCCTCTGCCCCACCTCTAGCGGCTCTTT	847
OY	781	gttccagttcctgtctctcc--tcaaaggcagccagagctgtgtcatgttctcatctc- -t	837
Db	848	GCTCCAGACTCTCCCTCCCTCTGTAGAGGCTGCTGGGCTGTTCACAGTGTTCATGCC	907
OY	838	-----acagagcatctcttgcctctctcttaacatccatccacacaactatccactc	891
Db	908	ACATAAATAACAATATTATTCACACTTATTATTCACACAAACCCACCGCCCACTCTCACACTC	967
OY	892	actagctcccaagccccctac-----ttaaccgtactcccccaccact	936
Db	968	ACTACTCTCCCACTCCCTGACCCCTTTGAGGCCCCCAAGTATCTGCAGACTCCCCCTGGCCA	1027
OY	937	cacccgaaccagtgatttacttgaacttltgtgac-----	968
Db	1028	CAGACCCCGAGGGCATTTGTTGTTCACTGTACTGTGGGCAAGATGGGTCACAGAGACC	1087
OY	969	-----caggacacttgatgaggtgtgacttgytgtgcaggaagccagaacccctggagctcg	1023
Db	1088	CACTTTCAGGCACTAAGAGGGGCTGTGACCTGTGGGAGGAAGCAAGAACACTGGGCTTAG	1147
OY	1024	gccagaagttcccaactgtgaaggggaagagctgtgaggaacactctctcctgtga-----tc	1079
Db	1148	GCCAGAGATGTCCCAATGTGAGGGGCGAGAACAA-GACACAGCTCTCTCCCTTGAGCAATTC	1206
OY	1080	cctgtgtgatttgaagaataactatttt	1108
Db	1207	CCTGTGATTTTAAACAGATATATTATTT	1235
RESULT	8		
LOCUS	AX180714	898 bp	DNA
DEFINITION	Sequence 1 from Patent WO0145730.	linear	PAT 06-AUG-2001
ACCESSION	AX180714		
VERSION	AX180714.1	GI:15132570	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Wiley,S.R.		
TITLE	Tweak receptor		
JOURNAL	Patent: WO 0145730-A 1 28-JUN-2001;		
	IMMUNEX CORPORATION (US)		


```

OY 418 gtaetgtcaaggtgacatttgatgaggaagaagctgttactgaagctgtagctgtgct 477
DB 833 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
OY 478 gaacgagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 537
DB 893 ggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
OY 538 tgggtcccaagctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 597
DB 953 cgggccccagctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1012
OY 598 ccttgagatccgcaacctcctcctgctgctgctgctgctgctgctgctgctgctgctgct 657
DB 1013 ccttcgagatccgcaacctcctcctgctgctgctgctgctgctgctgctgctgctgctgct 1072
OY 658 tggagcttttcaagtttcaagtttcaagtttcaagtttcaagtttcaagtttcaagtttca 717
DB 1073 cggagcttttcaagtttcaagtttcaagtttcaagtttcaagtttcaagtttcaagtttca 1132
OY 718 ctccagagagacacacacacacacacacacacacacacacacacacacacacacacacacac 776
DB 1133 ctcc-----cctgcacacacacacacacacacacacacacacacacacacacacacacac 1187
OY 777 ccttgctcagctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 834
DB 1188 cttgtgctcagctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1247
OY 835 tcc-----acagagatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 887
DB 1248 tccacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1307
OY 888 ccttcaagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 932
DB 1308 cctcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1367
OY 933 cactcaccgcaacacacacacacacacacacacacacacacacacacacacacacacacac 968
DB 1368 gccacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1427
OY 969 -----cagagacacacacacacacacacacacacacacacacacacacacacacacac 1019
DB 1428 acccacttccagcacttccagcacttccagcacttccagcacttccagcacttccagcacttcc 1487
OY 1020 ctgaagcagaagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1077
DB 1488 ctgaagcagaagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1546
OY 1078 ---tccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1130
DB 1547 atttcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1606
OY 1131 aaatgatatttaataataataataataataataataataataataataataataataataata 1159
DB 1607 aaatgatatttaataataataataataataataataataataataataataataataataata 1635

```

```

RESULT 10
AC016876/c 190358 bp DNA linear HTG 03-SEP-2001
LOCUS Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered
pieces.
AC016876 AC016876 GI:15421989
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 190358)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Howland,J., J. Gardyna,S., Grant,G., Hagos,B., Heath,A., Klein,J.,
Lehoczky,J., Liou,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL
COMMENT

```

```

REFERENCE
2 (bases 1 to 190358)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Howland,J., J. Gardyna,S., Grant,G., Hagos,B., Heath,A., Klein,J.,
Lehoczky,J., Liou,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL
COMMENT

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: submissions@genome.wi.mit.edu

----- Project Information -----
Center project name: 13849
Center clone name: 186_B_7

----- Summary Statistics -----
Sequencing vector: M13; M7815; 3% of reads
Sequencing vector: Plasmid; n/a; 97% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181928 bases at least Q40
Consensus quality: 18529 bases at least Q30
Consensus quality: 18735 bases at least Q20
Insert size: 16800; agarose-fp
Insert size: 18858; sum-of-coverage
Quality coverage: 8.1 in Q20 bases; agarose-fp
Quality coverage: 7.2 in Q20 bases.

* NOTE: This is a working draft sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1 20703: contig of 20703 bp in length
20704 20803: gap of 100 bp
20804 21935: contig of 1132 bp in length
21936 22035: gap of 100 bp
22036 24202: contig of 2167 bp in length
24203 24302: gap of 100 bp
24303 28199: contig of 3897 bp in length
28200 28299: gap of 100 bp
28300 35021: contig of 6722 bp in length
35022 35121: gap of 100 bp
35122 40745: contig of 5624 bp in length
40746 40845: gap of 100 bp
40846 47136: contig of 6291 bp in length
47137 47236: gap of 100 bp
47237 58980: contig of 11744 bp in length
58981 59080: gap of 100 bp
59081 68164: contig of 9084 bp in length
68165 68264: gap of 100 bp
68265 94625: contig of 26361 bp in length
94626 94725: gap of 100 bp
94726 102564: contig of 7839 bp in length
102565 102664: gap of 100 bp
102665 116764: contig of 14100 bp in length

```

```

* 116765 116864: gap of 100 bp
* 116865 136408: contig of 19544 bp in length
* 136409 136508: gap of 100 bp
* 136509 156380: contig of 19872 bp in length
* 156381 156480: gap of 100 bp
* 156481 180631: contig of 24151 bp in length
* 180632 180731: gap of 100 bp
* 180732 190358: contig of 9627 bp in length.
Location/Qualifiers
source
1. 190358
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-186B7"
/clone_11b="RP11-11 Human Male BAC"
1. 20703
/misc_feature
/note="assembly_fragment"
clone_end:5P6
vector_side:left"
misc_feature
20804..21935
/note="assembly_fragment"
misc_feature
22036..24202
/note="assembly_fragment"
misc_feature
24303..28199
/note="assembly_fragment"
misc_feature
28300..35021
/note="assembly_fragment"
misc_feature
35122..40745
/note="assembly_fragment"
misc_feature
40846..47136
/note="assembly_fragment"
misc_feature
47237..58980
/note="assembly_fragment"
misc_feature
59081..68164
/note="assembly_fragment"
misc_feature
68265..94625
/note="assembly_fragment"
misc_feature
94726..102564
/note="assembly_fragment"
misc_feature
102665..116764
/note="assembly_fragment"
misc_feature
116865..136408
/note="assembly_fragment"
misc_feature
136509..156380
/note="assembly_fragment"
misc_feature
156481..180631
/note="assembly_fragment"
misc_feature
180732..190358
/note="assembly_fragment"
clone_end:17
vector_side:right"
BASE COUNT 45173 a 48605 c 50635 g 44441 t 1504 others
ORIGIN
Query Match 26.0%; Score 303.8; DB 2; Length 190358;
Best Local Similarity 71.0%; Pred. No. 1.7e-68;
Matches 574; Conservative 0; Mismatches 162; Indels 73; Gaps 10;
QY 425 caggttcacattgaggaaggaaggtctctacctaagaactgctgtgtaacggt 484
DB 83053 CAGGTTCACATTGAGGAGGAAGGCTGCTACTACCTGAAGCTGACTGCTGGATGGT 82994
QY 485 gtgtgcgcctgcgtgcctgtgaagaattctcagcacacagcagcctctccctggcc 544
DB 82993 GTGTGCCTGCGCTGCGCTGCGGAGGAATTCTCAGCACACGCGGAGTTCCCTCGGGCC 82934
QY 545 cagcttcgcttgcagagtgctgtggtgcgtgcgcctgcgcagaggtctccctcgg 604
DB 82933 CAGCTTCGCTTGCAGAGTGCTGTGCTGTGCGCCCTGCGGCGCAGGTCCTCCCTGCGG 82874
QY 605 atccgacacctccctgggctctcttaaggctgccccctcttaacttaacttggacc 664
DB 82873 ATCCGACACCTCCCTGGCGCCATCTCAAGGCTGCCCTCTCTCACTACTTGGAGCTC 82814

```

```

QY 665 tttaagttcactgaggggagcctgtcctccagatctccttaacttccctgtgctccagg 724
DB 82813 TTCAAGTTCACTGAGAGGGCCCTGGTCTCCCGCAGTGCTCCAGAGCTCCGGCTCC--- 82755
QY 725 agcatcaccaactctccttaaccaccccaactcctcctcaacccctc-gctgtccttgt 783
DB 82756 --CTTCGACAGCTCTCTGCGGACCCCGTCCCTCTGCGCCACCTCCAGCGCTCTTTGCT 82699
QY 784 ccagtcctgtctctcc--tcaaggcagccagagctgtctacatgcttccatcc----- 837
DB 82698 CCAGACCTCCCTCCCTCCCTTAAGAGCTGCTGCGCTGTTCACGtGTTTTCATCCACA 82639
QY 838 ---acagacgtatcctgtcctctcttaacatccatccacacacacttccactcaact 894
DB 82638 TAAATACACTATTCCACACTTATGCTTACACTCCCCACCCACACCTCCACTCCACT 82579
QY 895 agctcccaagccctac-----tatacctgtactcccccacccactaac 939
DB 82578 ACTCTCCCAATCCCTGACCCCTTTGAGGCCCCAGTGATCTGACATCCCTCGGCGACAG 82519
QY 940 ccgaccagtgcttatgtactgtgac----- 968
DB 82518 ACCCCAGGGCATTTGTCTCACTGTACTCTGTGGCAGAGATGGTCCAGAAAGACCCAC 82459
QY 969 --caggcactgagatggtcgtgacctgtgycaggaagccagagacactgagactagcc 1026
DB 82458 TTCAGGACACTAAGAGGGCGCTGGACCTGGCGGAGAGCAAGAGACTGGGCTTAGGGC 82399
QY 1027 aqaagttcccaactgtgagggggaagagctgaggacaagctcctccctgga-----tccct 1082
DB 82398 AGAGTTCCTCCAAATGTGAGGGCGGAG--AACAAAGCAACACTCTCTCTTGAGAAATTCCT 82340
QY 1083 gtggatttgaagaa--agatacatttcttattatctgtgacaaaatgt---taaatga 1137
DB 82339 GTGGATTTTAAACAGATATTATTATTATTATTGTGACAAATGTTGATTAATGGA 82280
QY 1138 taataagagaataaatacatgattctct 1166
DB 82279 TATTAAATGAAATTAAGTACTAGCTCTCT 82251

RESULT 11
AC098923/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AC098923 179030 bp DNA linear HTG 20-DEC-2001
Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS
*** 68 unordered pieces.
AC098923.4 GI:17973852
HTG: HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 179030)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarta,J.,
Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileve,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burich,P., Burrell,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthalite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Barnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

```

Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Koryah, J., Kovar, C., Kratovic, J., Kureshi, R., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louieghed, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nkokenko, S., Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, R., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 179030)

Morley, K.C.

Direct Submission

Submitted (06-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:17064677.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G10K

Center clone name: CH230-154B15

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to findPhrapList

Consensus quality: 137096 bases at least Q40

Consensus quality: 145080 bases at least Q30

Consensus quality: 152950 bases at least Q20

Estimated insert size: 138980; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-IP estimation

Quality coverage: 2.2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length

* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank-draft-data.html>).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 68 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 7467: contig of 7467 bp in length

* 7468 7567: gap of unknown length

* 7568 12115: contig of 4548 bp in length

* 12116 12215: gap of unknown length

* 12216 20313: contig of 8098 bp in length

* 20314 20413: gap of unknown length

* 20414 25589: contig of 5176 bp in length

* 25590 25689: gap of unknown length

* 25690 31434: contig of 5745 bp in length

* 31435 31534: gap of unknown length

* 31535 37595: contig of 6061 bp in length

* 37596 37695: gap of unknown length

* 37696 42172: contig of 4477 bp in length

* 42173 42272: gap of unknown length

42273 46234: contig of 3962 bp in length

* 46235 46334: gap of unknown length

* 46335 50605: contig of 4271 bp in length

* 50606 50705: gap of unknown length

* 50706 55202: contig of 4497 bp in length

* 55203 55302: gap of unknown length

* 55303 58407: contig of 3105 bp in length

* 58408 58508: gap of unknown length

* 58509 61287: contig of 2779 bp in length

* 61288 61386: gap of unknown length

* 61387 64467: contig of 3081 bp in length

* 64468 64567: gap of unknown length

* 64568 68498: contig of 3931 bp in length

* 68499 68598: gap of unknown length

* 68599 73449: contig of 4851 bp in length

* 73450 73549: gap of unknown length

* 73550 77403: contig of 3853 bp in length

* 77403 77503: gap of unknown length

* 77503 81027: contig of 3525 bp in length

* 81028 81127: gap of unknown length

* 81128 84144: contig of 3017 bp in length

* 84145 84244: gap of unknown length

* 84245 86680: contig of 2436 bp in length

* 86681 86780: gap of unknown length

* 86781 89210: contig of 2430 bp in length

* 89211 89310: gap of unknown length

* 89311 92339: contig of 3029 bp in length

* 92340 92439: gap of unknown length

* 92440 94951: contig of 2512 bp in length

* 94952 95051: gap of unknown length

* 95052 95052: contig of 3535 bp in length

* 95053 98587: gap of unknown length

* 98587 98686: gap of unknown length

* 98687 100597: contig of 1911 bp in length

* 100598 100697: gap of unknown length

* 100698 102766: contig of 2069 bp in length

* 102767 102866: gap of unknown length

* 102867 105237: contig of 2371 bp in length

* 105238 105337: gap of unknown length

* 105338 107305: contig of 1968 bp in length

* 107306 107405: gap of unknown length

* 107406 109057: contig of 1652 bp in length

* 109058 109157: gap of unknown length

* 109158 111454: contig of 2297 bp in length

* 111455 111554: gap of unknown length

* 111555 113704: contig of 2150 bp in length

* 113705 113804: gap of unknown length

* 113805 115933: contig of 2129 bp in length

* 115934 116033: gap of unknown length

* 116034 118193: contig of 2160 bp in length

* 118194 118293: gap of unknown length

* 118294 120414: contig of 2121 bp in length

* 120415 120514: gap of unknown length

* 120515 122311: contig of 1797 bp in length

* 122312 122411: gap of unknown length

* 122412 125628: contig of 3217 bp in length

* 125629 125728: gap of unknown length

* 125729 128545: contig of 2817 bp in length

* 128546 128645: gap of unknown length

* 128646 130563: contig of 1918 bp in length

* 130564 130663: gap of unknown length

* 130664 132945: contig of 2282 bp in length

* 132946 133045: gap of unknown length

* 133046 134556: contig of 1511 bp in length

* 134557 134656: gap of unknown length

* 134657 136478: contig of 1822 bp in length

* 136479 136578: gap of unknown length

* 136579 138159: contig of 1581 bp in length

* 138160 138259: gap of unknown length

* 138260 140456: contig of 2197 bp in length

* 140457 140556: gap of unknown length

* 140557 142544: contig of 1988 bp in length

* 142545 142644: gap of unknown length

* 142645 143784: contig of 1140 bp in length

Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguni,M., Okwunonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuboan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,U., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vanson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 118395)
Worley,K.C.

Direct Submission
Submitted (07-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16756150.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GIJO
Center clone name: CH230-191M14
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 85618 bases at least Q40
Consensus quality: 94111 bases at least Q30
Consensus quality: 99937 bases at least Q20
Estimated insert size: 55807; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-IP estimation
Quality coverage: 0.7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
3372: contig of 3372 bp in length
3373
3472: gap of unknown length
3473
7021: contig of 3549 bp in length
7022
7121: gap of unknown length
7122
9925: contig of 2804 bp in length
9926
10025: gap of unknown length
10026
12351: contig of 2326 bp in length
12352
12451: gap of unknown length
12452
15703: contig of 3252 bp in length
15704
15803: gap of unknown length
15804
18388: contig of 2585 bp in length
18389
18488: gap of unknown length
18489
21644: contig of 3156 bp in length
21645
21744: gap of unknown length
21745
24057: contig of 2313 bp in length
24058
24157: gap of unknown length
24158
25917: contig of 1760 bp in length
25918
26017: gap of unknown length
26018
27985: contig of 1968 bp in length
27986
28085: gap of unknown length

28086
29145: contig of 1060 bp in length
29146
29245: gap of unknown length
29246
31057: contig of 1812 bp in length
31058
31157: gap of unknown length
31158
33961: contig of 2804 bp in length
33962
34061: gap of unknown length
33749: contig of 1688 bp in length
35750
35750: gap of unknown length
35850
37495: contig of 1646 bp in length
37496
37595: gap of unknown length
37596
39453: contig of 1858 bp in length
39454
41475: contig of 1922 bp in length
41476
41575: gap of unknown length
41576
43934: contig of 2359 bp in length
43935
44034: gap of unknown length
44035
45739: contig of 1705 bp in length
45740
45839: gap of unknown length
47091: contig of 1252 bp in length
47191: gap of unknown length
47192
48852: contig of 1661 bp in length
48853
48953: gap of unknown length
50055: contig of 1103 bp in length
50056
50155: gap of unknown length
51577: contig of 1422 bp in length
51578
51677: gap of unknown length
53662: contig of 1965 bp in length
53663
53762: gap of unknown length
53763
55525: contig of 1763 bp in length
55625: gap of unknown length
55626
56640: contig of 1015 bp in length
56641
56740: gap of unknown length
56741
58876: contig of 2136 bp in length
58877
58976: gap of unknown length
58977
60443: contig of 1467 bp in length
60444
60543: gap of unknown length
60544
61942: contig of 1399 bp in length
61943
62042: gap of unknown length
62043
63377: contig of 1335 bp in length
63378
63477: gap of unknown length
64892: contig of 1415 bp in length
64893
64992: gap of unknown length
64993
67305: contig of 2312 bp in length
67306
67404: gap of unknown length
67405
68639: contig of 1235 bp in length
68640
68740: gap of unknown length
68741
69861: contig of 1122 bp in length
69862
69961: gap of unknown length
71276: contig of 1315 bp in length
71277
71376: gap of unknown length
72603: contig of 1227 bp in length
72604
72703: gap of unknown length
72704
74279: contig of 1576 bp in length
74280
74379: gap of unknown length
74380
76194: contig of 1815 bp in length
76195
76294: gap of unknown length
76295
77477: contig of 1183 bp in length
77478
77578: gap of unknown length
77579
78686: contig of 1109 bp in length
78687
78786: gap of unknown length
80295: contig of 1509 bp in length
80296
80395: gap of unknown length
80396
81556: contig of 1161 bp in length
81557
81656: gap of unknown length
81657
82923: contig of 1267 bp in length
82924
83023: gap of unknown length
84800: contig of 1777 bp in length
84801
84900: gap of unknown length
84901
86054: contig of 1154 bp in length
86055
86154: gap of unknown length
86155
88128: contig of 1974 bp in length
88129
88228: gap of unknown length
89419: contig of 1191 bp in length

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 11:56:14 ; Search time 3297.36 Seconds
(Without alignments)
8713.678 Million cell updates/sec

Title: US-09-245-198a-3
Perfect score: 1373
Sequence: 1 atgcatgttagacttga.....gacaaatgtgataatgg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

1	1320.2	96.2	1353	6	AX201324
2	1320.2	96.2	1368	9	AF055872
3	1285	93.6	1306	9	AF030099
4	1236.4	89.3	1236	6	ARI40407
5	1096.8	79.9	1651	9	BC019047
6	768.2	56.0	190358	2	AC016876
7	629.2	45.8	898	6	AX180714
8	614.6	44.8	1168	10	AF030100
9	309.6	22.5	203083	2	AC069459
10	309.6	22.5	234182	10	AL603707
11	158.6	11.6	179030	2	AC098923
12	87.2	6.4	7218	6	166494
13	77.6	5.7	125020	9	AF429315
14	74.8	5.4	303091	2	AC084799
15	70.8	5.2	27571	2	AC105667
16	70.8	5.2	73948	2	AC022556
17	70.6	5.1	144979	2	AC016280
18	70.4	5.1	220469	2	AC074307
19	70.2	5.1	298166	2	AC087563
20	69.2	5.0	936	8	CNS01AYK
21	69.2	5.0	68631	2	AC094811
22	69.2	5.0	82897	2	AC022595
23	68.2	5.0	110737	2	AC011105
24	68.2	5.0	265537	2	AC087228
25	68	5.0	300695	2	AC079431
26	67.6	4.9	626	8	CNS018RP
27	67.2	4.9	66624	2	AC101520
28	67.2	4.9	147124	2	AC022768
29	67.2	4.9	177883	2	AC046159
30	67.2	4.9	181988	2	AC090552
31	67	4.9	74138	2	AC021272
32	67	4.9	158167	2	AC034105
33	66.8	4.9	62649	2	AC022552
34	66.8	4.9	75628	2	AC021793
35	66.8	4.9	100511	2	AC010774
36	66.4	4.8	840	8	CNS01BNS
37	66.2	4.8	30168	2	AC096156
38	66.2	4.8	61834	2	AC068105
39	66.2	4.8	79355	2	AC021280
40	66.2	4.8	82024	2	AC023210
41	66.2	4.8	150695	2	AC023576
42	66	4.8	66952	2	AC105944
43	66	4.8	89994	2	AC021263
44	66	4.8	134940	2	AC018939
45	66	4.8	137878	2	AP003300

ALIGNMENTS

RESULT 1
AX201324 LOCUS 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
AX201324 ACCESSION
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1353)
AUTHORS Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
Hillan,K.J., Marsters,S.A., Pan,J., Pitter,R.M., Roy,W.A., Smith,V.,
Stone,D.M., Watanabe,C.K. and Wood,W.I.
TITLES Compositions and methods for the treatment of tumour
JOURNAL Patent: WO 0153486-A 3 26-JUL-2001;
Genentech, Inc. (US)

FEATURES
source
location/Qualifiers
1..1353
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t

BASE COUNT 272 a 443 c 389 g 264 t
ORIGIN

Query Match 96.2%; Score 1320.2; DB 9; Length 1368;
Best Local Similarity 99.8%; Pred. No. 2.8e-220;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

49 cgaacccctcgagtcgagatgggagcggatgaagcagagacagagcccccacccatg 108
1 CGATCCCTCGGGTCCGGGATGGGGGGGGGTGAGGCGACGACGCCGCCGCCCATG 60

109 gccgcccctcgagagcagagagcagaggggagcgggagggagccagccctctg 168
61 GCCGCCGCTGGAGACCAAGAGCGGAGGGGGGGGGGGGAGCCGCCGCCCTCTG 120

169 gtccgcctcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgag 228
121 GTCCCGCTCGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 180

229 gtcaagtttgagagagcagagcagagtcgagtcgagtcgagtcgagtcgagtcg 288
181 GTCAAGTTTGGGGAGCCGGGGATCCGCTGCGCCAGAGCCTGCGCAAGAGAGCTGGTG 240

289 gcaagagagagcagagcagagccgctcgagagtcagatccacagagagagagatcc 348
241 GCAGAGAGAGAGCAGAGACCCGCTGGAGCTGATCCCGAGACAGAGAGAGAGAGATCCT 300

349 gcgccttcctgaacgagtcagtcgagtcgagtcgagtcgagtcgagtcgagtcg 408
301 GCGCCTTCTCGAACCAGCTAGTTGCGGCTCGCAGAAAGTGCACTAAAGCGGAGAAACA 360

409 cggagtcgagagagagagcagagccattatgaagttatccacagagtcgagtcgag 468
361 CGGCTCGAAG 420

469 gcgagagcagagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 528
421 GCCGAGCAGAGTCGTGGAGCGGAGACAGTAGTGCTGGGAGAGAGCAGAGATACAGCTCC 480

529 agcccttcgagtcagacccgagagtcgagtcgagtcgagtcgagtcgagtcgagtc 588
481 AGCCCTTGGGCTAACACCGCAGATGCGGAGATTTATAGCACCGGCGGCTCTAC 540

589 tacctgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 648
541 TACCTGAGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAGTCG 600

649 ctggtgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 708
601 CTGGTGGATGGTGCGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 660

709 tcctcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 768
661 TCCTCGGCGGCCCGCAGCTCCGCTCTGCGCAGGCTGTGGCGCTGGGCTGGGCTGGG 720

769 tccctcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 828
721 TCCCTCGGCGAGTCGCGACCTCCCTCGGCGCATCTCAAGAGCTGGCCCTCTCTCAC 780

829 tactcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 888
781 TACTTCGAGACTCTCCAGGTTCACTGAGGGGCCCTGGGCTCCCGGAGTGTCCAGGCT 840

889 gccgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 948
841 GCCGCGCTCCCTCGACAGCTCTCTGGGCGACCGGCTCCCTCTGCGCCACCTCGAGCGCT 900

949 ctggtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 1008
901 CTTTGGCTCGAGACTGCGCCCTCTCTAGAGAGCTGCTGGGCTGTTCACGTTGTTTCCA 960

1009 tccacataaatacagatattccacattatctacaaactcccccacgcgccaacttcca 1068

Db 961 TCCACATAAATACAGATATCCACACTGTATCTAGAACATCCCGCAGCGCCACTGTCCA 1020

Qy 1069 cctactagctccccaatccctgagcccttgagggcccccagtcagtcagtcagtcagtc 1128

Db 1021 CCTACTAGAGCTCCCAATCCCTGTGAGCTTGTAGGCCCCCATGTGATCTGATCCCTG 1080

Qy 1129 gccacagaccccccagagcatgtgttactgactctgtgtgagagagatgagtcagag 1188

Db 1081 GCCACAGACCCCGAGGTCATTTGTTTCACTGTACTCTGTGGCAAGATGGGTCCAGAG 1140

Qy 1189 acccactcagcagcactaagagggagtcgagccctgagcagagagagcagagtcgagc 1248

Db 1141 ACCCAGCTTACAGCAGTAAGAGGGGCTGAGACTGTGGCGGAGAGAGCCAGAGAGCTGG 1200

Qy 1249 ctgagcagagagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 1308

Db 1201 CTAGGCGAGAGGTTCCCAATGTGAGGGCGCGAAGACAGACAGCTCTCCCTTGAGAA 1260

Qy 1309 ttccctgagattttaaagagatattattatattatattatattatattatattatatt 1368

Db 1261 TTCCCTGTGATTTTAAACAGATATTATTATTATTATTATTATTATTATTATTATTATT 1320

Qy 1369 aatgg 1373

Db 1321 AATGG 1325

RESULT 3
AF030099
LOCUS AF030099 1306 bp mRNA linear PRI 20-DEC-1997
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1306)
Chicheportiche, Y., Bourdon, P., Xu, H., Hsu, Y.M., Scott, H.,
Hession, C., Garcia, I., and Browning, J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J Biol Chem. 272 (51), 32401-32410 (1997)
JOURNAL 98070415
MEDLINE 2 (bases 1 to 1306)
REFERENCE Bourdon, P., Hession, C., Tizard, R. and Browning, J.
ATTHORS Direct Submission
TITLE Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
JOURNAL Cambridge, MA 02142, USA
FEATURES
source
1. 1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue="tonsil"
18..767
/note="ligand in the TNF family; secreted protein; start
codon not verified experimentally"
/codon_start=1
/product="TWEAK"
/protein_id="AAC51923.1"
/translation="MAARRSRRRRGRGEPALVPLALGLIALACGLLAVSL
GSRASLSAEPQAELEVAEDDPSEINPQTESOPAPFLNVRPRRSAPGRKTR
ARRAIAHYEVRHPGODGAGVGVSGVSEARINSSPFLNVRNIGEIFYTRGL
VYLGVHDECKAAVYIKLIDVGLAIRCLIEFSAATASLSGLPOLRQCYSGLLAL
RESSSRIRLPLPAHLKAPFLYRFLFOVH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN

Db	361	GGAGCGAGGAGGTGTGAGCGGAGAGTGTGGGAGGAAGCAACAGTCAACAGC	420
Qy	526	ttcagccctctgagctacaaccgagatcgaggattatagtcacccggcggtc	585
Db	421	TCAGGCTCTGCGCTACAAACCGCAGATCGGAGATTATATAGTCACCCGGGCTG	480
Qy	586	tactaacctgactcagctgagctgacttgatagagggaaggtctgctactaacttaacgtgac	645
Db	481	TACTACTGTACTGTACAGTGTCACTTTGATGAGGGAAGCGTGTACTACTAAGCTGTGAC	540
Qy	646	ttgctgtgtgagtgatgtgtgtgcccctgcgtcgtctggaagaaattctcagcaactgtggcc	705
Db	541	TTGCTGTGTGATGTGTGTCTGGCCCTGGCGCTGTGAGAAATTCAGCCACTGCGCGC	600
Qy	706	agttccctcgaggcccaagctcgcctctgcaagtgatcctggtgtggccctcgagca	765
Db	601	AGTTCCCTCGGCCCCAGCTCCGCTCTGCGAGTGTGTGGCTGTGGCTTCGCGCA	660
Qy	766	gggtctccctctgagatcgcagccctccctgggcccatacacaagctgcccctctc	825
Db	661	GGGTCTCCCTCGGAGATCCGACCTCCCTGGGCGCATCTCAAGGCTGCCCTTCCTC	720
Qy	826	acctactcggactctctccaggtcactgaaggccctgtctcccaagtcgtccag	885
Db	721	ACCTACTTCGACACTCTTCAGGTTCACCTGAGGGGCGCTCTCCCGACATCGTCCAG	780
Qy	886	gtctcgagctccctctgaaagctctctgggcaaccgttccctctgcccacccctcagcc	945
Db	781	GCTCTTGCTTCACACTGCTGAGGAGTGTGGGAGCCGCTCTGCGCCACCTCAACC	840
Qy	946	gctcttctcagacactcgcctccctctctagaagctgtgctggcctgttcaagt	1005
Db	841	GCTCTTGCTTCACACTGCTGAGGAGTGTGGGAGTGTGGGAGTGTGACGTGTTT	900
Qy	1006	ccatccacataaatacagatattccactcttacttaacaactcccccaccccaactct	1065
Db	901	CCATCCCAACATTAATACAGTATTCCTCACTTATCTTACAACAACACCCACCCACTCT	960
Qy	1066	ccactactcagactccccaatcccttgaccccttgaggcccccagtgatctgactcccc	1125
Db	961	CCACTACTACTAGCTCCCAATCCCTGAGCCCTTGTAGGCCCCCAGCTGATCTGCCCTCC	1020
Qy	1126	ctggcacaagaccgccagggatgtgttctcactgtactctgtggcaagatgagtcag	1185
Db	1021	CTGGCCACAGACCCCGCAGGGGATGTGTCTACTCTGTGCGCAAGATGGGTCCAG	1080
Qy	1186	aagaccactcagcactaagaggcgtgagccctggcggcaggaagccaagaagactg	1245
Db	1081	AAGACCCCACTTCAAGCACTTAAGAGGGGCTGTGACCTGGCGCAGGAAGCCAAAGACTG	1140
Qy	1246	ggcctagggcaggaggttcccaatgtgaaggaggagaacaagaagaagctctcccttga	1305
Db	1141	GGCCTAGGCCAGAGATTCCTCAATGTAGGGGCGAAGAAAGCAAGACTCTCCCTTGA	1200
Qy	1306	gaattccctgtgattttttaaacaagatatattt 1341	
Db	1201	GAATTCCCTGTGATTTTAAACAGATATTATT 1236	
RESULT	5		
LOCUS	BC019047	1651 bp	mRNA
DEFINITION	BC019047	1651 bp	mRNA
ACCESSION	BC019047	1651 bp	mRNA
VERSION	BC019047.1	GI:17512138	
KEYWORDS	complete cds.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1651)		

AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Louis Staudt cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@hgti.nih.gov Shvachenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stanthip, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov</p> <p>Series: IRAL Plate: 30 Row: P Column: 5</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.</p> <p>location/Qualifiers</p> <p>1..1651</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="MGC:20669 IMAGE:4766071"</p> <p>/tissue_type="Primary B-Cells from Tonsils"</p> <p>/clone_id="N1H-MGC_48"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOTB7"</p> <p>106..510</p> <p>/codon_start=-1</p> <p>/product="Similar to tumor necrosis factor (ligand) superfamily, member 12"</p> <p>/protein_id="AAH19047.1"</p> <p>/db_xref="GI:17512139"</p> <p>/translation="MAARRSORRRGRGEGTLLVPLAIGLIALAGLLAVSLGSRASUSADPEADPELDOPSELNPDYESODPAPFLNRLVRRSPAPGRKTRARRAIAHYEVHRRPGDGAADAGGYTCRLRP"</p>
CDS	<p>106..510</p> <p>/codon_start=-1</p> <p>/product="Similar to tumor necrosis factor (ligand) superfamily, member 12"</p> <p>/protein_id="AAH19047.1"</p> <p>/db_xref="GI:17512139"</p> <p>/translation="MAARRSORRRGRGEGTLLVPLAIGLIALAGLLAVSLGSRASUSADPEADPELDOPSELNPDYESODPAPFLNRLVRRSPAPGRKTRARRAIAHYEVHRRPGDGAADAGGYTCRLRP"</p>
BASE COUNT	344 a 517 c 481 g 309 t
ORIGIN	
Query Match	79.9%; Score 1096.8; DB 9; Length 1651;
Best Local Similarity	84.8%; Pred. No. 2.2e-181;
Matches 1347; Conservative	0; Mismatches 2; Indels 239; Gaps 1;
Qy	25 tccgcccgcggctcccccctcccgatccctgggttcgggagtgagggtgagg 84
Db	25 TCCGCCCGCGGCTCCCTCCCTCCCGATCCCTCGGATGGGGGGCGGTGAGC 84
Qy	85 caggcacagcccccgcgccatagcgcccgctggaagccagagcgaggcgccgg 144
Db	85 CAGGCACAGCCCCCGCCCATGCGCCCGCTGGAGCCAGAGGGGGGCGCCGG 144
Qy	145 gggagccggcagccgcctgtgtcccgctcgcgtgagcctgagcctgtgctggcc 204
Db	145 GGGAGCGCGGACACGCGCTGTGTCGCCGCTCGCGTGGGCGTGGCGCTGGCGCC 204
Qy	205 tgctctggctcctcgtcgtggcggtgagttgtgggagccggagatgcttccggccag 264
Db	205 TGCTTGCGCTCTCTGCTGGCGGTGATGATTGGGAGCCCGGATGCTTCCGCCAG 264
Qy	265 gagcctgcccagggagcgtgtgtgcagagagagaccagcccgctcggaactgactcc 324

Db	265	GAGCTGCCCCGAGGAGGAGCTGTGTGGCAGAGAGGACAGGACCGGTCGGAACCTGAATCCC	324
Qy	325	cagacgaagaagaagccaaggaatcctctgcgccttcctcgaaecgaactatgtctcggcctcgca	384
Db	325	CAGACAGAAAGAAACCCGAGGATCCTGGCCCTTTCGTAACCGCACTAGTTGCGGCTCGCAGA	384
Qy	385	agtcacactaaagcccgaaacaacagagcttcgaagaagcatcgaccccaataabaact	444
Db	385	AGTCACACTTAAGGCCCGGAAACACGGGCTCGAAGAGCGCATGCGACCCATTATGAACATT	444
Qy	445	catccaagaecttgaaacaagaagcgcaggg-----	475
Db	445	CATCCAGCACTTGAAACGACGAGCAGGACGACGAGATGAGGTTACACAACATTGCTGAGG	504
Qy	476	-----	475
Db	505	CCATGAGTACTAAGTGGTGGACCAAGATTGAAACCCACGCTAGAAATGTGCTGTACT	564
Qy	476	-----	475
Db	565	GGACATGGTCTGCATGAAAGCGCAGGNGGCTGTGACAGGGGTGAAGGGTTCATCGAGGG	624
Qy	476	-----	475
Db	625	CCACATCCAAAAAGGGAGAGGAAGTTTCCAGAAAGAAAGAGACACATCTCCACCA	684
Qy	476	-----caagtctgagcagggaaatgaaatgagctgg	505
Db	685	TTACCAAGAGGTTCAAAGAGAACAGACAGACAGTGTGACCGGACAGTGAATGGCTGG	744
Qy	506	aggaagccagaatcaaacagctccagccctctgcgtatacaaacgcgcagaatcgaggaaatta	565
Db	745	AGGAAGCCAGATATCAACAGCTCCAGGCCCTTGCGCTACCAACCGCAGATCGGGGAGTTTA	804
Qy	566	taagcaaccggagctgggctcttaactactgacgtcaagctgaactttatagaaggggaag	625
Db	805	TAGTACACCCGGGCTGGGCTTACTACTGTACTGTAGGTGACATTGTATGAGGGGAAGG	864
Qy	626	ctgtctaacctgaagctggaactctgctgtgaaatgtgtctgagccctgcgtactgaag	685
Db	865	CTGTCTACCTGAACCTGGAATCTGTGGTGAATGTGTGCTGGCCCTGGGCTCTCTGAGG	924
Qy	686	aattctaaagcaactgcygcgaagtccctcggggcccaagctccgcgtcttcgcaagtgctgt	745
Db	925	AATTCTGACCACTGCGGCGCAAGTTCCCTGGGCCCCAGCTCCGCACTGCGCAGAGGTGTG	984
Qy	746	gscgtgttgcccttcgagccagaggttcctcctctgcgaatccgcaaccccttcgagccatc	805
Db	985	GGCTTTGGCCCTCGGCGCAAGGGTCGCTCGCGAATCGGCAACCTCCCTCGGCCCATC	1044
Qy	806	lcaagagctgcgcccttcctcaactaacttcggaactcttccaagttcaactgaagggccctg	865
Db	1045	TCGAAGGCGCCCTTCTCACTACTGTGGACTTTCACAGGTTTCACTGAGGGGCCCTGG	1104
Qy	866	ctccccaacatctgctccagagctgcgcgctcccccctgcgaagctctctggaaccgcgtcc	925
Db	1105	TCTCCCGCAATCGTCCAGGCTGCGGGCTCCCTTGACAGCTCTGTGGGACCCGGTCC	1164
Qy	926	ccctctgcccaacccttcagacgcgtctctgtctcgaagcctgcgcctccctctagaagctgc	985
Db	1165	CCCTGCCCCCAACCTTCAGCGGCTTTTGTGTCAGACACTCCCTCCCTCTTAAGAGCTGCC	1224
Qy	986	ctgggcctgttaagtggttttccatcccaaaataaatacaatctccacactatctatca	1045
Db	1225	TGGGCTGTTCACGTGTTTTCATCCACATAATACAGATTTCACACTTATATTGACA	1284
Qy	1046	actccccaaccgcgaactctcaactcaactcaagatccccaatccctgaaccctcttgagccc	1105
Db	1285	ACTCCCAACGCGCCACTGTCCACACTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCC	1344
Qy	1106	ccaagtatctgaatccccctctggccaagaaccgccagggcaatgtgttcaactatctct	1165

Db	1345	COAGGATCTGACATCCCCCTGGCCACAGACCCCAAGGAGATTGTCTCACTGACTCT	1404
Oy	1166	gtggcgaagatggatggatccagaagagaccacttcctgaagcctaagagggtctgacctggcg	1225
Db	1405	gtggcgaagatggatggatccagaagagaccctcactgagcactaaagggctggacctggcgc	1464
Oy	1226	gaagaaacccaagaagatggagccttagagccagagatccccaatctgaaggaggcagaagac	1285
Db	1465	GCAGGAACCCAAAGAGAGATGGGGCTTAGGCCAGAGATTCGCCAAATGTGAGGGGGCGAANAAC	1524
Oy	1286	aagacaagctccctcccttgagaatccctctgagatctttaaacaagatcatcttattc	1345
Db	1525	AAGACAACCTCCTCCTCTTGAGAAATTCCTGTGGAATTTTAAACAGAAATTAATTTTAAT	1584
Oy	1346	attatctgcacaaatgtgataatgg	1373
Db	1585	ATTATTGTGACAAATGTTGATTAATGG	1612
RESULT	6		
AC016876/c			
LOCUS	AC016876	190358 bp	DNA
DEFINITION	Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered		
ACCESSION	AC016876		
VERSION	AC016876.4	GI:15421989	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 190358)		
TITLE	Birren,B., Linton,L., Nusbaum,C. and Lander,E.		
JOURNAL	Homo sapiens chromosome, clone RP11-186B7		
REFERENCE	2 (bases 1 to 190358)		
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,D., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donelan,D., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,D., Gardyna,S., Grant,G., Hagos,B., Harford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoccky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,D., Morrow,J., Naylor,D., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,D., Teffaye,S., Tirelli,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
COMMENT	On Sep 3, 2001 this sequence version replaced gi:13431059. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
	Genome Center		
	Center: Whitehead Institute/ MIT Center for Genome Research		
	Center code: WIBR		
	Web site: http://www-seq.wi.mit.edu		
	Contact: sequence_submissions@genome.wi.mit.edu		
	Project Information		
	Center project name: L3849		
	Center clone name: 186_B_7		
	Summary Statistics		
	Sequencing vector: M13; M7815; 3% of reads		
	Sequencing vector: Plasmid; n/a; 97% of reads		
	Chemistry: Dye-terminator Big Dye; 100% of reads		
	Assembly program: Phrap; version 0.960731		
	Consensus quality: 181928 bases at least 40		
	Consensus quality: 185529 bases at least 30		
	Consensus quality: 187335 bases at least 20		


```
RESULT 7
AX180714
LOCUS AX180714 898 bp DNA Linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 898
location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
52. .873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC5004.1"
/db_xref="GI:15132571"
/translation="MATGSRISLLAFGLLPLWDEGSATSDRMKQIEDKIEILS
KIYHIEIARIKILIGERTSSLSRASLSAQEPAGEELVAEEDPDESLNPTEES
ODPAFLRLVRRPSAKGRKTRARIAAHYEHPRGDDAGADGVSGWEAA
RINSSPLRYRQICEFTYRAGLYLYLCOVHFDGKRVYIAKDLVGVYALRLTE
FSATRASLSGLPOLRCOVSGLLALPSSLSLRILPMAHLAAAPLTYFGLEOVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Query Match 45.8%; Score 629.2; DB 6; Length 898;
Best Local Similarity 99.5%; Pred. No. 5.9e-100;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 232 agttggagagcggagcagcgtgtctgcgcgcagagagcctgccagagagcgtgtgtgca 291
DB 250 aatttggagagcggagcagcgtgtctgcgcgcagagagcctgccagagagcgtgtgtgca 309
QY 292 gaggagagcagagaccgtctcgaaactgaatcccccagagagagagcagatcctcg 351
DB 310 gagagagcagagaccgtctcgaaactgaatcccccagagagagagcagatcctcg 369
QY 352 ccttcctcgaacagactatctgcgcgcagagagagcagcctcgaagagcagagagagc 411
DB 370 ccttctcgaacagactatctgcgcgcagagagagcagcctcgaagagagcagagagc 429
QY 412 gtcgagagagcagcagcagccatctagatctcaccagcagcctgcagagagagcag 471
DB 430 gctcgagagagcagcagcagccatctagatctcaccagcagcctgcagagagagcag 489
QY 472 caggcagagtgtagcagagcagtgtagcgtggaagagagagagagagagagagcagc 531
DB 490 caggcagagtgtagcagagcagtgtagcgtggaagagagagagagagagagagcagc 549
QY 532 cctctgcgctaaacagcagagatcgaggaggtttatagttacccgggctgtgctctac 591
DB 550 cctctgcgctaaacagcagagatcgaggaggtttatagttacccgggctgtgctctac 609
QY 592 cgtacatcgacagtgcaactttagagagagagagcgtctacatcgagagagagagagc 651
DB 610 cgtacatcgacagtgcaactttagagagagagagcgtctacatcgagagagagagagc 669
QY 652 gtgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 711
DB 670 gtgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 729
QY 712 ctcgagagagagagagagagagagagagagagagagagagagagagagagagagagag 771
DB 730 ctcgagagagagagagagagagagagagagagagagagagagagagagagagagagag 789
```

```
QY 772 tccctgcgagatccgcagccctccctgggagccatctcaagagcgtccctctcaccctac 831
DB 790 tccctgcgagatccgcagccctccctgggagccatctcaagagcgtccctctcaccctac 849
QY 832 ttcgagactctccaggttcaactgtagagggccctgg 865
DB 850 ttcgagactctccaggttcaactgtagagggccctgg 883

RESULT 8
AF030100
LOCUS AF030100 1168 bp mRNA Linear ROD 20-DEC-1997
DEFINITION Mus musculus TWEAK mRNA, partial cds.
ACCESSION AF030100
VERSION AF030100.1 GI:2707220
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
location/Qualifiers
1. 1168
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="peritoneal macrophages"
<1. .679
/note="ligand in the TNF family; secreted protein"
/codon_start=2
/product="TWEAK"
/protein_id="AAC53517.1"
/db_xref="GI:2707221"
/translation="VLSGLALACILGLLVVSLGSWATLSAQPSEELTAEDRREP
PELNQTEESODVPEFLQVLRPSAKGRARRAIAAHYEHVHPRGDDAGV
DGVSGWEETKINSPLRYRQICEFTYRAGLYLYLCOVHFDGKRVYIAKDLV
GVYALRLTEERSATRASLSGLPOLRCOVSGLLALPSSLSLRILPMAHLAAAPLTY
FGLEOVH"
BASE COUNT 242 a 360 c 298 g 268 t
ORIGIN

Query Match 44.8%; Score 614.6; DB 10; Length 1168;
Best Local Similarity 75.8%; Pred. No. 2e-97;
Matches 906; Conservative 0; Mismatches 219; Indels 70; Gaps 9;

QY 178 ggcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgc 237
DB 2 gtcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgcgtgcgc 61
QY 238 gggagcggagcagcgtctgcgcgcagagagcctgcgcagagagagagagagagagag 297
DB 62 gggagcggagcagcgtctgcgcgcagagagcctgcgcagagagagagagagagagag 121
QY 298 gaccagagcagcgtcggaactgaatcccccagagagagagagagagagagagagagag 357
DB 122 cgcggagagcagcgtcggaactgaatcccccagagagagagagagagagagagagagag 181
QY 358 ctgaaccgagctagtcgagcagagagagagagagagagagagagagagagagagagagag 417
DB 182 ttggagacactagtcgagcagagagagagagagagagagagagagagagagagagagag 241
```


OY	418	agaagagatcagagcccatatgaatctatccacagacccggagacggaacgaagga	477
Db	242	CGAGTATTGACAGCCCATTTATGAGTTCACTTCGACCAGGACAGAGATGACACAGCA	301
OY	478	ggtctgagacggaacgctagctgctcggagaggaagccagatacaacagctccagccctctg	537
Db	302	GGTGTGGATTGGGACAGTGAAGTGGTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG	361
OY	538	cgcatacaacgcagatcggggagtttatagtacccgggctcgggctctactactgttac	597
Db	362	CGCTACGACACCGCAATTGGGGAAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC	421
OY	598	tgtcagtgagacattgatagaggggaagcgtgctctaccctgaagctgagacttgttgytat	657
Db	422	TGTCAAGTGACATTTGATGATGAGGAAAGCTGTCTACCTGAACCTGGACTGTGTGGTAAC	481
OY	658	ggtgtcgtcggccctgcgctcgtctgtgaggaattctcaagcaacgcgagccagttccctcgag	717
Db	482	GGTGTGCTGGCCCTCGCGCTCGCTGGAGAAATTTCTCAGCCACAGCAAGCAAGCTCTCTGGG	541
OY	718	ccccagctccgcctctgtccaaagtgtctcgtgctgttggccctcggcgcgaagttccctctg	777
Db	542	CCCCAGCTCCGTTTGTGCGACAGGTGCTGTGGGCTGTTGGCCGCTCGGCGCAAGGCTTCCCTT	601
OY	778	cggatccgacacccctccctgggcccattcaagagctgcgcccctctctcaactacttgya	837
Db	602	CGGATCCGACCCCTCCCTGGGCTCATTTAAAGGTGGCCCTCTTCAACTTATCTTTGGA	661
OY	838	ctcttcacaggttcaactgaggggcccgtgctctcccaacagtcgtcccaagctcgcggtcc	897
Db	662	CTCTTTCAAGTTCACTGAGGGGCTTGGCTGCCAGATTCCCTTAACCTTCCCTGGCTCC	721
OY	898	-----ctctgacagctctcttgggaacccggctccctctgcgccacccctcaagcgcgtcttt	952
Db	722	AGGAGCATACACACACATCCCTACCCCAACCCCACTCCCTCCACACCCCTC-GCTGGCTCTT	780
OY	953	gcttcagagcgtccctctccctctctagaagctgctgctggcgtgtgttaagtggttttccatcc	1012
Db	781	GGTCCAGTCTCTGTCTCTC--TCAAAGAGGACCAAGCTTGTTCACATGTATTCACATTC--	837
OY	1013	acataataacaglatctccactctatcttacaactcccccacgcgcacccctccacctc	1072
Db	838	-----ACAGACGATTCCTGTGCTCTTCTTAATCCCATCCACCAACACTATCCACCTC	891
OY	1073	actagctccccaatccctcgtgacccctttgagccccagatgactcgaactcccccttgcga	1132
Db	892	ACTAGCTCCCAAAAGCCCTAC-----TTATCCCTGACTCCCCACCCCACT	936
OY	1133	cagacccccagagcttgtgttctactgtactctgtggygaagaaatggygtccagaagacc	1192
Db	937	CACCCGACACCTGTTTATTGACTTTGTGCAC-----	968
OY	1193	cacttcagagcaactaagagggtctgacctgcgcgaggaagccaaagagactggtcgtag	1252
Db	969	-----CAGGCACTGAGATGGGCTGTGACCTCGGTGCGCAGAGACCAAGAACTGTGGGACTG	1023
OY	1253	gccagagattcccaaatgtgaggggcgagaaacaa-gacaagctctccctccttggaaattc	1311
Db	1024	GCCAAAGATTTCCCAACTGCTGAGGGGGAAGAGCTGGGGGACAACTCCTCCCTGGA---TC	1079
OY	1312	cctgtgagttttaaacaagatatatttttatattatattgtgacaaaatgtgtga	1366
Db	1080	CTGTGGATTGTTGAAA--AGATCTATTATTTATTTATTTATGTGACAAAATGTTAA	1132

RESULT	9
AC069459/C	-
LOCUS	203083 bp DNA linear HTG_27-JUN-2001
DEFINITION	Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
SEQUENCE	, 7 unordered pieces.
ACCESSION	AC069459
VERSION	AC069459..23 GI:14547768
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS-FULLTOP.

SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 230383)
Metzker, M. L., Lewis, L. R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okunowo, G., Carlino, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonini, D., Brooks, A., Brown, J., Butley, C., Bunick, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcun-Tansey, J., Gill, R., Gorrell, J. H., Gunaratne, P., Haller, C., Hernandez, J., Hognes, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Koval, C., Liu, J., Liu, W., Loulseged, H., Lozado, R. J., Martin, R., Massey, E., McLeod, M. P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B., Perez, L., Reller, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Murthy, D. M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G., Worley, K. and Gibbs, R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 203083)
Worley, K. C.
TITLE Direct Submission
AUTHORS Submitted (31-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL On Jun 25, 2001 this sequence version replaced gi:12621364.
COMMENT ----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: MAFO
Center clone name: RP23-168P5
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 5% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210656; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 62152: contig of 62152 bp in length
* 62153 62252: gap of unknown length
* 62253 118772: contig of 56520 bp in length
* 118773 118872: gap of unknown length
* 118873 148924: contig of 30052 bp in length
* 148925 149024: gap of unknown length
* 149025 167231: contig of 18207 bp in length
* 167232 167331: gap of unknown length
* 167332 189907: contig of 22576 bp in length
* 189908 190008: gap of unknown length
* 190008 196537: contig of 6530 bp in length
* 196538 196637: gap of unknown length
* 196638 203083: contig of 6446 bp in length.
Location/Qualifiers

61177	63386:	gap of unknown length
61387	64636:	contig of 3081 bp in length
64468	64567:	gap of unknown length
64568	66498:	contig of 3931 bp in length
68499	66598:	gap of unknown length
68599	72449:	contig of 4851 bp in length
73450	72549:	gap of unknown length
73550	77402:	contig of 3853 bp in length
77403	77502:	gap of unknown length
77503	81027:	contig of 3525 bp in length
81028	81127:	gap of unknown length
81128	84144:	contig of 3017 bp in length
84145	84244:	gap of unknown length
84245	86680:	contig of 2436 bp in length
86681	86780:	gap of unknown length
86781	89210:	contig of 2430 bp in length
89211	89310:	gap of unknown length
89311	92339:	contig of 3029 bp in length
92340	92439:	gap of unknown length
92440	94951:	contig of 2512 bp in length
94952	95051:	gap of unknown length
95052	98586:	contig of 3535 bp in length
98587	98686:	gap of unknown length
98687	100597:	contig of 1911 bp in length
100598	100697:	gap of unknown length
100698	102166:	contig of 2069 bp in length
102167	102866:	gap of unknown length
102867	105337:	contig of 2371 bp in length
105238	105237:	gap of unknown length
105338	107305:	contig of 1968 bp in length
107306	107405:	gap of unknown length
107406	109057:	contig of 1652 bp in length
109058	109157:	gap of unknown length
109158	111454:	contig of 2297 bp in length
111455	111554:	gap of unknown length
111555	113704:	contig of 2150 bp in length
113705	113804:	gap of unknown length
113805	115933:	contig of 2129 bp in length
115934	116033:	gap of unknown length
116034	118193:	contig of 2160 bp in length
118194	118293:	gap of unknown length
118294	120414:	contig of 2121 bp in length
120415	120514:	gap of unknown length
120515	122111:	contig of 1797 bp in length
122112	122411:	gap of unknown length
122412	125628:	contig of 3217 bp in length
125629	125728:	gap of unknown length
125729	128545:	contig of 2817 bp in length
128546	128645:	gap of unknown length
128646	130563:	contig of 1918 bp in length
130564	130693:	gap of unknown length
130694	132945:	contig of 2282 bp in length
132946	133045:	gap of unknown length
133046	134556:	contig of 1511 bp in length
134557	134656:	gap of unknown length
134657	136478:	contig of 1822 bp in length
136479	136578:	gap of unknown length
136579	138159:	contig of 1581 bp in length
138160	138259:	gap of unknown length
138260	140456:	contig of 2197 bp in length
140457	140556:	gap of unknown length
140557	142544:	contig of 1988 bp in length
142545	142644:	gap of unknown length
142645	143784:	contig of 1140 bp in length
143785	143884:	gap of unknown length
143885	145304:	contig of 1420 bp in length
145305	145494:	gap of unknown length
145495	147291:	contig of 1987 bp in length
147332	147491:	gap of unknown length
147492	148908:	contig of 1417 bp in length
148909	149008:	gap of unknown length
149009	150147:	contig of 1139 bp in length
150148	150247:	gap of unknown length

[illegible]

Db 1158 yy 1217
Qy 923 tccccctgccccccctcagccgctctgtccccagacctgcctccctcctcagagcct 982
Db 1218 yy 1277
Qy 983 gctcggcgctgttcacgtgttccatccacataaatacagatcccatctatcct 1042
Db 1278 yy 1337
Qy 1043 acaactccccaccgccactctccactcactagctccccaatccctgaaccttgag 1102
Db 1338 yy 1397
Qy 1103 cccccagtgatctcgaactccccctggcccaacccccaggcattgtgtcactgt 1160
Db 1398 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyygtaacaaattcttctatctct 1455

RESULT 13
LOCUS AF429315 125020 bp DNA linear PRI 18-JAN-2002
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
JOURNAL
MEDLINE 21583737
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Title Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNAL
FEATURES
source location/Qualifiers
1..125020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
Disease-like 2 (HDL2)"
complement(35381..35746)
/rpt_type=tandem
/rpt_unit=ctg
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/note="JPH3"
/note="JPH3"
complement(<36507..36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AAL0941.1"
/db_xref="GI:17646245"
/translation="MSSSGRFFPDGSGYCGGMEQCKAHGVCYCPKCGEYTGSM
HGFVGLVTPSGNTDGTMAQGRHGIGLSKGKWMYKGEVTHGFKRGYVRECA
NGAKYEGTMSNLDGCGTETSDG"
BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others

ORIGIN
Query Match 5.7%; Score 77.6; DB 9; Length 125020;
Best Local Similarity 11.0%; Pred. No. 0.00044;
Matches 96; Conservative 380; Mismatches 396; Indels 1; Gaps 1;
Qy 462 gacagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 521
Db 16871 GCARRKMKSHARRRRCITSKSMWSMBMSVSYVKHSHHABSCMHMKCTWSCCM 16930
Qy 522 cagctcagccctctgtcctacacacccagacccgagagcagcagcagcagcagc 581
Db 16931 MMYKSSMWGSMGCMCGRRRSKGMKMYSHGMSRSMRMTGSKMMRSSMCTSSCYAS 16990
Qy 582 gctctactactgtactgtcagcagcagcagcagcagcagcagcagcagcagcagc 641
Db 16991 MCMCMGSMCCCMRSCCMRSCCMRSCCMRSCCMRSCCMRSCCMRSCCMRSCCM 17050
Qy 642 ggaactgtcgtgtgagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 701
Db 17051 SRMCSRSRSKCKSRGCGSGMGKGGKSTGRKTRKSKRMGAKMVMTYRRSMRKM 17110
Qy 702 ggcacgttccctc-gggcccaagctccgcctctgcagagtgctgtgtgtgtgtgt 760
Db 17111 YSSKGMVCMYCMCGRRRCGCTCSCMTSRSAMCCSYCAKCKSMYSCYGMKMGYIYACS 17170
Qy 761 ggcacaggtctccctctgcagcagcagcagcagcagcagcagcagcagcagcagc 820
Db 17171 YRSMSRSKYCMRSGSTYSTGCCCTTTTCCCCNANTGCGGAAGCTTTTNCMKTYSY 17230
Qy 821 tctctacactactctgcagcagcagcagcagcagcagcagcagcagcagcagcagc 880
Db 17231 RKRNGACMKYNNYNNNSMSSCRAGMSCTKYKSSMTMSMASYCMWCMYTCMRSMAS 17290
Qy 881 cccaagctgcgcgcctccctcagcagcagcagcagcagcagcagcagcagcagcagc 940
Db 17291 GMSMSYMKMKMSMRNSCMYCMWCMKMCYCMWMSMRNSMGMYMAYMSKSSRDCYCT 17350
Qy 941 cagcgcctctgtctcagcagcagcagcagcagcagcagcagcagcagcagcagc 1000
Db 17351 RCYWCMSKSKSYKSYMMMSRKRMKMSRCMSMGMSASRSCYKCYKSMKSSMMSK 17410
Qy 1001 gtttcacatccacataaatacagatctccactctatcttcaactcccccagccgc 1060
Db 17411 CYACAGCMKMKGYMKYMKRMSKRRMWSKRSMTRRMRMKGAGAMMKCTSRMSRCA 17470
Qy 1061 actctcagctcactagctccccaatccctgaccttgagcgcacagcagcagcagc 1120
Db 17471 MMKMSYCASRSCAMMSGTYMKCASYCMMSYCTCMTSYMYRCTCMCKGMSYTKS 17530
Qy 1121 cccccctgcagcagacaccccaagcagcagcagcagcagcagcagcagcagcagcagc 1180
Db 17531 KSMSSMSYKGRKRKRSYCMCTSRGAMSCWRRCYKRGASSMRAGSSRRARRKSG 17590
Qy 1181 tccagaagaccacactcagcagcagcagcagcagcagcagcagcagcagcagcagc 1240
Db 17591 RSKWMMTGMMRSKTYTCTGRPMAMTYMCCMRRRRMTYRRMSAMGMSKSSWSGMRM 17650
Qy 1241 gactggcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1300
Db 17651 GSASSRRCASRSMSCMSRSMKMGMSRSMSCMSKMMGSGRSASAKCKGSGMRRRSR 17710
Qy 1301 ctgagaatccctgtgagatctttaaacagat 1333
Db 17711 KYRRKRGCKKRSMTKSGSKGSKCWMKRSRSGMT 17743

RESULT 14
AC084799/c 303091 bp DNA linear HTG 17-NOV-2000
LOCUS AC084799/c
DEFINITION Mus musculus chromosome 16 clone RP23-197M9, WORKING DRAFT
SEQUENCE, 101 unordered pieces.

ACCESSION AC0084799
VERSION AC0084799.1 GI:11192127
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
2 (bases 1 to 303091)
REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 0
Center clone name: RPCL-23_197M9

Summary Statistics
Consensus quality: 152568 bases at least Q40
Consensus quality: 175579 bases at least Q30
Consensus quality: 186949 bases at least Q20
Estimated insert size: 198300; agarose-fp estimation
Estimated insert size: 293091; sum-of-contigs estimation
Quality coverage: 4.85 in Q20 bases; agarose-fp estimation
Quality coverage: 3.28 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 101 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1117 1116: contig of 1116 bp in length
1217 1216: gap of unknown length
2537 2536: contig of 1320 bp in length
2637 2636: gap of unknown length
4058 4057: contig of 1421 bp in length
4158 4157: gap of unknown length
5246 5245: contig of 1088 bp in length
5346 5345: gap of unknown length
6462 6461: contig of 1116 bp in length
6562 6561: gap of unknown length
7726 7725: contig of 1164 bp in length
7826 7825: gap of unknown length
9144 9143: contig of 1318 bp in length
9244 9243: gap of unknown length
10493 10492: contig of 1249 bp in length
10593 10592: gap of unknown length
11667 11666: contig of 1074 bp in length
11767 11766: gap of unknown length
12951 12950: contig of 1184 bp in length
13051 13050: gap of unknown length
14286 14285: contig of 1235 bp in length
14386 14385: gap of unknown length
15479 15478: contig of 1093 bp in length
15579 15578: gap of unknown length
16780 16779: contig of 1201 bp in length
16880 16879: gap of unknown length
18279 18278: contig of 1399 bp in length
18379 18378: gap of unknown length
19453 19452: contig of 1074 bp in length
19553 19552: gap of unknown length
21013 21012: contig of 1460 bp in length
21112 21112: gap of unknown length

21113 21169: contig of 1057 bp in length
22270 22269: gap of unknown length
22270 23367: contig of 1098 bp in length
23368 23467: gap of unknown length
23468 24559: contig of 1092 bp in length
24560 24659: gap of unknown length
24660 25825: contig of 1166 bp in length
25826 25925: gap of unknown length
25926 27036: contig of 1111 bp in length
27037 27136: gap of unknown length
27137 28245: contig of 1109 bp in length
28246 28345: gap of unknown length
28346 29636: contig of 1311 bp in length
29637 29756: gap of unknown length
29757 30883: contig of 1127 bp in length
30884 30983: gap of unknown length
30984 32125: contig of 1142 bp in length
32126 32225: gap of unknown length
32226 33397: contig of 1172 bp in length
33398 33497: gap of unknown length
33498 34645: contig of 1148 bp in length
34646 34745: gap of unknown length
34746 35894: contig of 1149 bp in length
35895 35994: gap of unknown length
35995 37156: contig of 1162 bp in length
37157 37256: gap of unknown length
37257 38396: contig of 1140 bp in length
38397 38496: gap of unknown length
38497 39570: contig of 1074 bp in length
39571 39670: gap of unknown length
39671 40852: contig of 1182 bp in length
40853 40952: gap of unknown length
40953 42116: contig of 1164 bp in length
42117 42216: gap of unknown length
42217 43372: contig of 1156 bp in length
43373 43472: gap of unknown length
43473 44603: contig of 1131 bp in length
44604 44703: gap of unknown length
44704 45924: contig of 1221 bp in length
45925 46024: gap of unknown length
46025 47174: contig of 1150 bp in length
47175 47274: gap of unknown length
47275 48744: contig of 1470 bp in length
48745 48844: gap of unknown length
48845 49969: contig of 1125 bp in length
49970 50069: gap of unknown length
50070 51135: contig of 1066 bp in length
51136 51235: gap of unknown length
51236 52690: contig of 1455 bp in length
52691 52790: gap of unknown length
52791 53920: contig of 1130 bp in length
53921 54020: gap of unknown length
54021 55101: contig of 1081 bp in length
55102 55201: gap of unknown length
55202 56324: contig of 1123 bp in length
56325 56424: gap of unknown length
56425 57583: contig of 1159 bp in length
57584 57683: gap of unknown length
57684 58822: contig of 1139 bp in length
58823 58922: gap of unknown length
58923 60547: contig of 1625 bp in length
60548 60647: gap of unknown length
60648 61743: contig of 1096 bp in length
61744 62998: contig of 1155 bp in length
62999 63098: gap of unknown length
63099 64220: contig of 1122 bp in length
64221 64320: gap of unknown length
64321 65398: contig of 1078 bp in length
65399 65498: gap of unknown length
65499 66762: contig of 1264 bp in length
66763 66862: gap of unknown length
66863 68025: contig of 1163 bp in length

*	68026	gap of unknown length
*	68126	68125: config of 1275 bp in length
*	68126	68126: config of 1275 bp in length
*	69401	69500: gap of unknown length
*	69501	70718: config of 1218 bp in length
*	70719	70818: gap of unknown length
*	70819	71995: config of 1177 bp in length
*	71996	72095: gap of unknown length
*	72096	73322: config of 1137 bp in length
*	73233	73352: gap of unknown length
*	73333	74476: config of 1144 bp in length
*	74477	74576: gap of unknown length
*	74577	75863: config of 1287 bp in length
*	75864	75963: gap of unknown length
*	75964	77021: config of 1058 bp in length
*	77022	77121: gap of unknown length
*	77122	78212: config of 1091 bp in length
*	78213	78312: gap of unknown length
*	78313	79424: config of 1112 bp in length
*	79425	79524: gap of unknown length
*	79525	80617: config of 1093 bp in length
*	80618	80717: gap of unknown length
*	80718	81804: config of 1087 bp in length
*	81805	81904: gap of unknown length
*	81905	83007: config of 1103 bp in length
*	83008	83107: gap of unknown length
*	83108	84187: config of 1060 bp in length
*	84168	84267: gap of unknown length
*	84268	85347: config of 1080 bp in length
*	85348	85447: gap of unknown length
*	85448	86521: config of 1074 bp in length
*	86522	86621: gap of unknown length
*	86622	87744: config of 1123 bp in length
*	87745	87844: gap of unknown length
*	87845	88965: config of 1121 bp in length
*	88966	89065: gap of unknown length
*	89066	90176: config of 1111 bp in length
*	90177	90276: gap of unknown length
*	90277	91279: config of 1003 bp in length
*	91280	91379: gap of unknown length
*	91380	92563: config of 1184 bp in length
*	92564	92663: gap of unknown length
*	92664	93688: config of 1025 bp in length
*	93689	93788: gap of unknown length
*	93789	94934: config of 1146 bp in length
*	94935	95034: gap of unknown length
*	95035	96111: config of 1077 bp in length
*	96112	96211: gap of unknown length
*	96212	98352: config of 2141 bp in length
*	98353	98452: gap of unknown length
*	98453	99645: config of 1193 bp in length
*	99646	99745: gap of unknown length

[illegible][illegible]

RESULT	15
AC105667	
LOCUS	27571 bp DNA linear HTG 09-JAN-2002
DEFINITION	Rattus norvegicus clone CH230-12F12, *** SEQUENCING IN PROGRESS
ACCESSION	AC105667
VERSION	AC105667.1 GI:18092890
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

REFERENCE
AUTHORS

1 (bases 1 to 27571)

Muzny, D.M., Adams, C., Adio-Oduola, B., Allosman, F.R., Allen, C., Alstbrooks, S.L., Amaralunge, H.C., Are, J.R., Banks, T., Barbarta, J., Benton, J., Blinage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhey, C., Burch, P., Butrell, C., Butrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., DeThorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulys, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S., Joudah, S., Kalisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kralovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loutsengh, H., Lozardo, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mayhew, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Milner, G., Mitter, Z., Mitchell, T., Monabhat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nickerson, E., Nnokenko, S., Ogun, M., Okunodu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojts, A., Rojoudkan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Slisson, I., Sodergren, E., Sonakke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, C., Tellro, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watling, K., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Welstock, G. and Gibbs, R.

TITLE
JOURNAL
REFERENCE

2 (bases 1 to 27571)

Worley, K.C.

TITLE
JOURNAL
REFERENCE

Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

WMENT

```

Assembly program: Phrap; version 0.990329first call to
findphraplist
Consensus quality: 21300 bases at least Q40
Consensus quality: 25175 bases at least Q30
Consensus quality: 27665 bases at least Q20
Estimated insert size: 6068: sum-of-contigs estimation
Quality coverage: 0x 1n Q20 bases; agarose-fp estimation
Quality coverage: 0.1x 1n Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	2066:	contlg of 2068 bp in length
*	2069	2168:	gap of unknown length
*	2169	3384:	contlg of 1116 bp in length
*	3285	3384:	gap of unknown length
*	3385	4420:	contlg of 1036 bp in length
*	4421	4520:	gap of unknown length
*	4521	5587:	contlg of 1067 bp in length
*	5588	5687:	gap of unknown length
*	5688	6960:	contlg of 1273 bp in length
*	6061	7060:	gap of unknown length
*	7061	8243:	contlg of 1183 bp in length
*	8243	8343:	gap of unknown length
*	8344	9855:	contlg of 1512 bp in length
*	9855	9955:	gap of unknown length
*	9956	11027:	contlg of 1072 bp in length
*	11028	11272:	gap of unknown length
*	11128	12236:	contlg of 1109 bp in length
*	12337	12336:	gap of unknown length
*	12337	13771:	contlg of 1435 bp in length
*	13772	13871:	gap of unknown length
*	13872	15016:	contlg of 1145 bp in length
*	15017	15116:	gap of unknown length
*	15117	16143:	contlg of 1027 bp in length
*	16144	16243:	gap of unknown length
*	16244	17324:	contlg of 1081 bp in length
*	17325	17424:	gap of unknown length
*	17425	18674:	contlg of 1250 bp in length
*	18675	18774:	gap of unknown length
*	18775	20534:	contlg of 1760 bp in length
*	20535	20634:	gap of unknown length
*	20635	21800:	contlg of 1166 bp in length
*	21801	21900:	gap of unknown length
*	21901	23483:	contlg of 1583 bp in length
*	23484	23583:	gap of unknown length
*	23584	24753:	contlg of 1170 bp in length
*	24754	24853:	gap of unknown length
*	24854	26413:	contlg of 1560 bp in length
*	26414	26513:	gap of unknown length
*	26514	27371:	contlg of 1038 bp in length

FEATURES	location/Qualifiers
source	1. .27571
	/organism="Rattus norvegicus"
	/dbxref="taxon:10116"
	/clone="CH230-12F12"
BASE COUNT	5806 a 6848 c 7657 g 5287 t 1973 others
ORIGIN	

Query Match	5.2%	Score 70.8;	DB 2;	Length 27571;
Best Local Similarity	48.1%	Pred. No. 0.0075;		
Matches 201; Conservative	0;	Mismatches 217;	Indels 0;	Gaps 0;

QY	718	ccccagctcgcagctctcgcagaggtgcctcgtctgttcggccctgagggccaggtctccctcgtg	777
Db	24774	CGCCGCGCCCCCCCCCTCCGCGCGCTTCGCGCTTCCGCAACGCGCGCGCTCTCCGCCCC	25033
QY	778	cgagatccgaacctccctcgtggtccatctcaagctgcgccctctccatccactcttcgga	837
Db	25034	TGCCCCCTCACAACCCCCCCCCCCCCCCCCCCCCCGCTTCGCCCTCTCCCGCTCCCG	25095
QY	838	ctcttcacagttcaactagagggacctgtgtctccacacagttcgtcccaagctgcgcgtcc	897
Db	25094	CTCCTCCTCTCTCCCTCCGCCCGCGCCCGCCCGCCCCCCCCCCCCCCCCCGCGCTCC	25153
QY	898	cctcgcagctctctctcgtggagaccgcggtctccctctcgtccacacctcaagcagctcttctcc	957
Db	25154	CGCCGCGCCCCCGCCCCCTCCGCCCGCCCCCCCCCTCCCTCCCTCCGCCCGCCCCCGCCC	25213
QY	958	agacatcgccctccctcctctcagagctgcgtcctcgttcagctgtttctcatccacata	1017

Db 25214 CCCCCGCGCCCCCTCCGCCCGCCCTTCGCCCTTCCTCCGCCGCC 25273
QY 1018 aatacaglatitccactcttatcttacaactcccccaacgccaactctcaactcactag 1077
Db 25274 CCCCCGACCCCTCCGCCGCCGCCGCCGCCCTCCGCCGCCGCCCTCCACGCGCCCC 25333
QY 1078 ctcccaatcccttgacccttgaggccccaagtgatctcgactcccccttgccacag 1135
Db 25334 CCCCCCCCCCCCCCGGACCACTCGACACCTCCCTCCACGGGCTCGCGCGG 25391

Search completed: June 22, 2002, 12:11:08
Job time: 19703 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 11:53:24 ; Search time 328.09 Seconds
(Without alignments)
7184.990 Million cell updates/sec

Title: US-09-245-198a-3
1373
Sequence: 1 atgcattgttgacttga.....gacaaatgtgataatg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1373	100.0	1373	19	AAV18600
2	1325	96.5	1421	20	AAV56000
3	1320.2	96.2	1353	21	AAV49717
4	1226.4	89.3	1236	19	AAV47613
5	1226.4	89.3	1236	22	AAV4350
6	833.4	60.7	1030	20	AAV23424
7	629.2	45.8	898	22	AAV30964
8	614.6	44.8	1168	19	AAV18599
9	519.2	37.8	701	20	AAV23425

10	245.4	17.9	282	16	AAV22190	Human gene signal
C 11	70.4	5.1	114955	20	AAV53491	Human adenovirus A1
C 12	61	4.4	1337	20	AAV17263	Human gene express
C 13	60.6	4.4	1000	21	AAV02484	Human colon cancer
C 14	59.2	4.3	10732	21	AAV10594	Gene encoding a su
C 15	58	4.2	1218	21	AAV02488	Human colon cancer
C 16	54.4	4.0	1593	21	AAV02504	Human colon cancer
C 17	53.2	3.9	53526	19	AAV94101	Human PKD1 gene.
18	53.2	3.9	53577	17	AAV18551	Human polyomavirus
19	53.2	3.9	53577	19	AAV94108	Human PKD1 locus b
20	50	3.6	50	20	AAV56002	Human tumour necro
21	50	3.6	50	21	AAV49732	Human PRO207 DNA p
22	50	3.6	201	8	AAV70195	Streptomyces prote
23	50	3.6	2188	20	AAV7506	Human ovarian tumo
24	49.6	3.6	1517	21	AAV2028	CDNA encoding huma
25	49.2	3.6	825	20	AAV25378	Human mFLINT #2 nu
26	49.2	3.6	825	21	AAV51078	Human mature FLINT
27	49.2	3.6	936	20	AAV25376	Human FLINT #2 nuc
28	49.2	3.6	936	21	AAV8731	Human FAS ligand 1
29	49.2	3.6	936	21	AAV51076	Human FLINT coding
30	49.2	3.6	936	22	AAV07385	Human fas ligand 1
31	49.2	3.6	936	22	AAV8920	Nucleotide sequenc
32	49.2	3.6	936	22	AAV84738	Nucleotide sequenc
33	49.2	3.6	936	22	AAV84739	Nucleotide sequenc
34	49.2	3.6	936	22	AAV7696	Human FLINT native
C 35	49	3.6	1126	21	AAV02538	Human colon cancer
C 36	48.8	3.6	201	8	AAV70194	Signal portion of
C 37	48	3.5	320	21	AAV38185	Primer used in the
C 38	48	3.5	109519	22	AAV08693	Micromonospora DNA
39	48	3.5	114955	20	AAV53491	Human adenovirus A1
40	47.6	3.5	767	20	AAV23419	Human APO6 DNA. H
41	47.6	3.5	813	20	AAV25377	Human mFLINT #1 nu
42	47.6	3.5	813	21	AAV75999	DNA encoding a mat
43	47.6	3.5	813	21	AAV88730	Human FAS ligand 1
44	47.6	3.5	813	21	AAV51077	Mature human FLINT
45	47.6	3.5	813	22	AAV07380	Human mature fas 1

ALIGNMENTS

RESULT 1	
AAV18600	
ID AAV18600 standard; CDNA; 1373 BP.	
XX	
AC AAV18600;	
XX	
DT 21-JUN-1998 (first entry)	
XX	
DE Homo sapiens tumour necrosis factor related ligand (TNFRL) gene.	
XX	
KW TNFRL; tumour necrosis factor related ligand; tnfr; treatment;	
KW cancer; autoimmune disease; immune system; stimulation; suppression;	
KW graft rejection; ds.	
XX	
OS Homo sapiens.	
XX	
FT	Location/Qualifiers
FT CDS	1..852
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
PN	
W09805783-A1.	
XX	
PD 12-FEB-1998.	
XX	
PE 07-AUG-1997; 97WO-US13945.	
XX	
PR 18-MAR-1997; 97US-0040820.	
PR 07-AUG-1996; 96US-0023541.	
PR 18-OCT-1996; 96US-0028515.	
XX	
PA (BIOJ) BIOGEN INC.	

OS	Human sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	92..841
FT		/tag= a
FT		/product= "Apo-3 ligand"
XX		
PN	W09919490-A1.	
XX		
PD	22-APR-1999.	
XX		
PF	09-OCT-1998;	98WO-US21407.
XX		
PR	17-DEC-1997;	97US-0069862.
PR	10-OCT-1997;	97US-0062037.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Ashkenazi AJ, Marsters SA, Pittl R;	
XX		
RR	WPI: 1999-287982/24.	
RR	P-PSDB: AAY09369.	
XX		
PT	New human Apo-3 ligand (a tumour necrosis factor) homologue	
XX		
PS	Clatam 18; Fig 1; 74pp; English.	
XX		
CC	The present sequence encodes a human tumour necrosis factor (TNF) and	
CC	lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has	
CC	cytostatic activity. Apo-3 ligand can be used to induce apoptosis in	
CC	mammalian cancer cells, to induce NF-kappaB-dependent transcription and	
CC	to induce JNK/SAPK-dependent responses in mammalian cells.	
XX		
SO	Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;	

[illegible]

Db	483	gtcgtctggaggaagccagaatacaagctccagccctctgcgtctacaacgcgagatcg	542
Oy	557	gggaattatagtcaccgggctgggtctactaacctgaactctgaagggaatttgatg	616
Db	543	gggaattatagtcaccgggctgggtctactaacctgaactctgaagggaatttgatg	602
Oy	617	aggaggaagcgtctactagaagcttggaactctgctgtgtagtgatgtgctgcgtgcgt	676
Db	603	aggaggaagcgtctactagaagcttggaactctgctgtgtagtgatgtgctgcgtgcgt	662
Oy	677	gctcgagaggaattcttcagccaactcgcgcagattccctcgggcccagctccgcctgc	736
Db	663	gctctgagaggaattcttcagccaactcgcgcagattccctcgggcccagctccgcctgc	722
Oy	737	aggtctctggagctctttgagccctcgggcgaagggtccctccctggagatccgacccctcc	796
Db	723	aggtctctggagctctttgagccctcgggcgaagggtccctccctggagatccgacccctcc	782
Oy	797	gggcccattctcaagcgtgcgcctctctctcaacctactctggaactcttcagagttcaatg	856
Db	783	gggcccattctcaagcgtgcgcctctctctcaacctactctggaactcttcagagttcaatg	842
Oy	857	gggcccgtgtctccccaagtcgtctccaggtcgcgcgtctccctcgcagacgtctctgggc	916
Db	843	gggcccgtgtctccccaagtcgtctccaggtcgcgcgtctccctcgcagacgtctctgggc	902
Oy	917	accgcgtccctctgcgcgcacccctcaagcgcctctttgtctcagaacctgtgcctccctca	976
Db	903	accgcgtccctctgcgcgcacccctcaagcgcctctttgtctcagaacctgtgcctccctca	962
Oy	977	gagcgtcgtcgtgcgtctgtcaacgtgtttccatccccaataaatacagatattccactct	1036
Db	963	gagcgtcgtcgtgcgtctgtcaacgtgtttccatccccaataaatacagatattccactct	1022
Oy	1037	tatcttacaactcccccacgcgcacatctccacctcaactagctctcccaatccctgacct	1096
Db	1023	tatcttacaactcccccacgcgcacatctccacctcaactagctctcccaatccctgacct	1082
Oy	1097	ttgagagccccagatgactctgacccccctctggccaagaagcccccaaggcatttgattca	1155
Db	1083	ttgagagccccagatgactctgacccccctctggccaagaagcccccaaggcatttgattca	1141
Oy	1157	ctgtacctctgtggccaagatggtgtccagaagaagcccccaactctcaggaactaagaaggctc	1216
Db	1143	ctgtacctctgtggccaagatggtgtccagaagaagcccccaactctcaggaactaagaaggctc	1202
Oy	1217	gacctgctgcgcagagaagccaagaagactctggcctagggccaagagttcccaaatgtgaagg	1277
Db	1203	gacctgctgcgcagagaagccaagaagactctggcctagggccaagagttcccaaatgtgaagg	1263
Oy	1277	ggcgagaaacaagaacaagctctctccctctggaatctccctgtggaatttttaacaagatatt	1333
Db	1263	ggcgagaaacaagaacaagctctctccctctggaatctccctgtggaatttttaacaagatatt	1322
Oy	1337	attttattatattctgtgacaaaatgtgtgtaaatg	1373
Db	1323	attttattatattctgtgacaaaatgtgtgtaaatg	1359
RESULT 3			
ID AAA49717			
AAA49717 standard; cDNA, 1353 BP.			
XX AAA49717;			
XX AC			
XX XX			
XX 25-SEP-2000 (first entry)			
XX DE			
Human PRO207 cDNA clone DNA30879-1152.			
PRO207; human; antitumour; tumour; therapy; cytostatic;			
KM breast cancer; ovarian cancer; renal cancer; colorectal cancer;			
KM uterine cancer; prostate cancer;lung cancer; bladder cancer;			

KW		central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FH	CDS	/tag= a 58..807
FT	sigspeptide	b /*tag= 178..804
FT	matspeptide	c /*tag= c
XX		
PN	M0200037638-AZ.	
PD	29-JUN-2000.	
PF	02-DEC-1999;	99WO-US28565.
XX		
PR	22-DEC-1998;	98US-O113296.
PR	08-MAR-1999;	99WO-US05028.
PR	21-APR-1999;	99US-O130232.
PR	28-APR-1999;	99US-O131445.
PR	14-MAY-1999;	99US-O134287.
PR	20-JUL-1998;	99US-O144758.
PR	26-JUL-1999;	99US-O145698.
PR	15-SEP-1999;	99WO-US21590.
PR	15-SEP-1999;	99WO-US21547.
PA	(GENE) GENENTECH INC.	
XX	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA,	
PI	Napier MA, Piltl RM, Wood WI;	
DR	-P-PSDB: AAU95338.	
XX	WFI: 2000-442668/38.	
PT	Noel composition to inhibit neoplastic cell growth or for treating tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 -	
PS	Claim 20; Fig 3; 172pp; English.	
CC	The present sequence is that of cDNA clone DNA30879-1152 (ATCC 209358) encoding human PRO207 (see AAU95338), which shows homology to several members of the tumour necrosis factor family, especially human lymphotoxin (23.4%). The cDNA was identified in a foetal kidney cDNA library following identification of an expressed sequence tag with homology to human Apo-2 ligand. A claimed method for inhibiting the growth of a tumour cell comprising exposing the tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see AAU95337-49), their agonists or chimeric polypeptides incorporating them. The tumour is especially a cancer selected from breast, ovarian, renal, colorectal, uterine, prostate, lung, bladder and central nervous system cancer, melanoma and leukaemia. Nucleic acids encoding PRO179 etc. are used in the recombinant production of the antitumour polypeptides.	
SO	Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:	
OY	Query Match Best Local Similarity Score 96.2%; DB 21; Length 1353; Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Db	1 cgatccctcgagtcgccggaatggagggtgcacagcccccccatttg 60	
Oy	gcgcgcgatcggagaccagaagcgcaaggcgccggggagccgcgacttctg 168	
Dn	gcgcgcgatcggagaccagaagcgcaaggcgccggggagccgcgacttctg 120	

[illegible]

OY	1249	ctaggccagggagttccccaatggtgagggcgagagaacaagaagctctcccttgagaa	1308
Db	1201	cttaggcccagggagttccccaatgtgagggcgagaaacaagaagctctcccttgagaa	1260
OY	1309	ttccctgtggagtttttaaacagatattttttattttattgtgacaaatgttgata	1368
Db	1261	ttccctgtggagtttttaaacagatattttttattttattgtgacaaatgttgata	1320
OY	1369	aatgg 1373	
Db	1321	aatgg 1325	
RESULT 4			
AAV47613	ID	AAV47613 standard; cDNA; 1236 BP.	
XX	AAV47613;		
XX	AAV47613;		
DT	27-OCT-1998	(first entry)	
XX	TNF related endothelium proliferative agent gene.		
XX	ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;		
KW	tissue grafting; vascularisation; apoptosis; autoimmune; birth control.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	1..750	
FT	CDS	/*tag= a	
FT		/product= "TREPA"	
XX			
PN	WO9835061-A2.		
XX	13-AUG-1998.		
XX	12-FEB-1998;	98WO-US02859.	
PF	10-FEB-1998;	98US-0021706.	
PR	12-FEB-1997;	97US-0798692.	
XX	(ABBO) ABBOTT LAB.		
PA			
PI	Wiley SR;		
XX	WPI: 1998-447255/38.		
DR	P-PSDB; AAM29745.		
XX	Detecting nucleic acid encoding TREPA - useful for diagnosis and		
XX	treatment of autoimmune disease, tumours and inflammation		
XX	Claim 11; Page 123-4; 142pp; English.		
XX	The TNF-related endothelium proliferative agent (TREPA), or its		
CC	activators or agonists, are used to treat a deficit of TREPA, e.g. to		
CC	promote wound healing or tissue grafting, by promoting vascularisation,		
CC	also to induce apoptosis for treating cancer and eliminating autoreactive		
CC	T cells, as an adjunct to cancer chemotherapy or antiviral treatment.		
CC	TREPA peptides can also be used to target cytotoxic agents or for		
CC	affinity isolation of the corresponding receptor, the nucleic acid for		
CC	which can be used to transform tumour cells to render them more		
CC	responsive to TREPA and to screen for TREPA mimics.		
CC	Ribozymes, antisense RNA, antibodies or peptides, are used to treat		
CC	TREPA-associated diseases, e.g. tumours and metastases (by inhibiting		
CC	vascularisation), inflammation or a wide range of autoimmune conditions		
CC	conditions involving abnormal stimulation of epithelial cells (e.g.		
CC	atherosclerosis), for birth control (inhibiting ovulation and placental		
CC	formation) or other angiogenic conditions (e.g. ulcers).		
XX	Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;		

Query Match	89.3%	Score 1226.4	DB 19	Length 1236	
Best Local Similarity	99.5%	Pred. No. 8.4e-252			
Matches 1230	Conservative 0	Mismatches 6	Indels 0	Gaps 0	
QY	106	atgagccgcccgtccgagaccagagagcgagagagggcgccgagggagagccgagccgcccgtg	165		
Db	1	atggccgcgcccgtccgagaccagagagcgagagggcgcccgaggggagagccggagccgcccgtg	60		
QY	166	ctgtgccgcccctgcgcccctgtgcctgtgcggccctgcggccctgtgcctgtgcctgtgcgc	225		
Db	61	ctgtgccgcccctgcgcccctgtgcctgtgcggccctgcggccctgtgcctgtgcctgtgcgc	120		
QY	226	gtgtcactgtttggggagccgggcatctgcctgcgcgcccagagagctgtccagagagagctg	285		
Db	121	gtgtcactgtttggggagccgggcatctgcctgcgcgcccagagagctgtccagagagagctg	180		
QY	286	gtggcagagagagagaccagagaccggtccggaactcgaatccccacagaaagaaagccagagt	345		
Db	181	gtggcagagagagagaccagagaccggtccggaactcgaatccccacagaaagaaagccagagt	240		
QY	346	ccctgcgccccttcctgaaccgaactagtctgcgctccgacagagtgacactaaagccgagaa	405		
Db	241	ccctgcgccccttcctgaaccgaactagtctgcgctccgacagagtgacactaaagccgagaa	300		
QY	406	acacgaggtccgaagaagcgatcgcgaagcccatltaagaattcatccacagacctgtgacagagc	465		
Db	301	acacgaggtccgaagaagcgatcgcgaagcccatltaagaattcatccacagacctgtgacagagc	360		
QY	466	ggagcgccagagcgaggtgtgtgacgggacagtgaaatgtgtctgggagaaagccagatccaagc	525		
Db	361	ggagcgccagagcgaggtgtgtgacgggacagtgaaatgtgtctgggagaaagccagatccaagc	420		
QY	526	tccagccccttcgctgacaaaccgacagatccggggaggttaataagtcacccgggctgggctc	585		
Db	421	tccagccccttcgctgacaaaccgacagatccggggaggttaataagtcacccgggctgggctc	480		
QY	586	tactaacctgtactgtcaagtgacatttgaagaaggagaaagcgctctaccctgaagctgagac	645		
Db	481	tactaacctgtactgtcaagtgacatttgaagaaggagaaagcgctctaccctgaagctgagac	540		
QY	646	ttgcgtgggtgagatgtgtgtctgtgcctgtgcgctcgaggaaatttcagcaactgcggcc	705		
Db	541	ttgcgtgggtgagatgtgtgtctgtgcctgtgcgctcgaggaaatttcagcaactgcggcg	600		
QY	706	agttccctcgggcccacagctccgcccctgcgacaggtgtctcgggctgtgtgcctgcggcca	765		
Db	601	agttccctcgggcccacagctccgcccctgcgacaggtgtctcgggctgtgtgcctgcggcca	660		
QY	766	gggtccctccctgcgagatccgcaaccctccctctgaagccatccaagctgccccctctctc	825		
Db	661	gggtccctccctgcgagatccgcaaccctccctctgaagccatccaagctgccccctctctc	720		
QY	826	acctactccggaactcttcagagttcaactgagaggccctgtgtctcccaacagtctgccag	885		
Db	721	acctactccggaactcttcagagttcaactgagaggccctgtgtctcccaacagtctgccag	780		
QY	886	gttgcgcggtcccccctcgaaagatctctcttgagaaaccgggtccctctgcgcccacccaagcc	945		
Db	781	gttgcgcggtcccccctcgaaagatctctcttgagaaaccgggtccctctgcgcccacccaagcc	840		
QY	946	gtctcttgcacagacctgtgcctccctctctctctctctctctctctctctctctctctctct	1005		
Db	841	gtctcttgcacagacctgtgcctccctctctctctctctctctctctctctctctctctctct	900		
QY	1006	ccatcccaataataatacagtatctccaactcttatcttacaactcccccaacgccaactct	1065		
Db	901	ccatcccaataataatacagtatctccaactcttatcttacaacaccccacgccaactct	960		
QY	1066	ccactctcaatctcccaactcccgaaactcttgagccccccagagatctgaatcccccc	1125		
Db	961	ccactctcaatctcccaactcccgaaactcttgagccccccagagatctgaatcccccc	1020		
QY	1126	ctggccacagagaccccccagggcattgtgtcactatgtactctgtcgggcaagaaaggttccag	1185		

Db	1201	ctggccacagagaccccccaaggcatctgtgtccactctgtacacctctgtggcgaagtgtgtccag	1080
Qy	1186	aagaagcccaacttaaggaactaagaagggctgtgaacctgtgcgcgaaggaagccaaagagactg	1245
Db	1081	aagagccccaactctaaaggacataaagaagggctctgtgaacctgtgcgcgaaggaagccaaagagactg	1140
Qy	1246	ggccaaagcagcagaagatctcccaaaatcgtgaagggcgcgaagaacagaacaaagctcctcccttga	1305
Db	1141	ggcctagcgcaagagatctcccaaaatcgtgaagggcgcgaagaacagaacaaagctcctcccttga	1200
Qy	1306	gaattccctgtgaactctttaaaacagatatctttt	1341
Db	1201	gaattccctgtgaactctttaaaacagatatctttt	1236

(ABBO) ABBOTT LAB.
 Willey SR:
 XX WPI; 2001-280760/29.
 DR P-PSDB; AAE00891.
 XX
 PT Inducing angiogenesis in mammal at desired sites for promoting wound
 PT healing, by administering soluble fragment of extracellular domain of
 PT tumor necrosis factor related endothelium proliferative agent protein
 PT -
 XX
 PS Example 2; Column 73-74; 53pp; English.
 XX
 CC The present invention relates to extracellular signal molecules,
 CC particularly members of tumour necrosis factor (TNF) family molecules
 CC designated as TREPA (TNF related endothelium proliferative agent).
 CC Soluble biologically active TREPA are used to treat TREPA-associated
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
 CC in human for promoting wound healing and for vascularising grafted
 CC tissue for successful grafting and to promote tissue grafts.
 CC The present sequence is a cDNA clone ID #690050 encoding human TREPA.
 XX
 XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

[illegible]

|||||
Db 1021 ctggccaaagaccgccagggatctgtctactctgtggcaaggttggtccag 1080
Qy 1186 aagaccctctcagcactaagaggggctgacctgagcgcaagaaccagaagactg 1245
Db 1081 aagaccctctcagcactaagaggggctgacctgagcgcaagaaccagaagactg 1140
Qy 1246 ggcctagagccagaggttcccaatctgtgagggcgagaacaagaacgtctcccttga 1305
Db 1141 ggcctagagccagaggttcccaatctgtgagggcgagaacaagaacgtctcccttga 1200
Qy 1306 gaattccctgtgattttaaacaagatatttt 1341
Db 1201 gaattccctgtgattttaaacaagatatttt 1236
RESULT 6
AXX23424
ID AXX23424 standard; DNA; 1030 BP.
XX AXX23424:
DT 18-JUN-1999 (first entry)
XX Human TNRL3 DNA.
DE Tumour necrosis factor receptor: signal transducer molecule; TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KM apoptosis; human; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..627
FT /tag= a
FT /product= "TNRL3"
XX
PN MO9911791-A2.
XX 11-MAR-1999.
XX
PF 04-SEP-1998; 98MO-US181393.
XX
PR 05-SEP-1997; 97US-0924634.
XX
PA (UNIW) UNIV WASHINGTON.
XX Chaudhary PM.
XX
DR WPI: 1999-205191/17.
XX P-PSDB; AAW93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Example VII; Fig 13A; 156bp; English.
XX
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/ active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or

CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:

Query Match 60.7%; Score 833.4; DB 20; Length 1030;
Best Local Similarity 99.9%; Pred. No. 3.7e-168;
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 229 gtcaagttggagagccgggacatcgctgcgcccagagagcctgcgcaagagagctggtg 288
Db 1 gtcaagttggagagccgggacatcgctgcgcccagagagcctgcgcaagagagctggtg 60
Qy 289 gcagagagagagccagagccgctcggaatgaatcccgagacagagagagagagatcct 348
Db 61 gcagagagagagccagagccgctcggaatgaatcccgagacagagagagagagatcct 120
Qy 349 gcgccttcctgaaccgcagatgtcggcctcgcaagaatgcacctaagagcgcaaaaca 408
Db 121 gcgccttcctgaaccgcagatgtcggcctcgcaagaatgcacctaagagcgcaaaaca 180
Qy 409 cgggctcgaagagcgatgcagcccatatgaatgaltcatcacaagaccttgcagagagca 468
Db 181 cgggctcgaagagcgatgcagcccatatgaatgaltcatcacaagaccttgcagagagca 240
Qy 469 gcgcagagagagtgtagagagagagagtgtagtgcttgcggagagagagcagagatacaagctcc 528
Db 241 gcgcagagagagtgtagagagagagagtgtagtgcttgcggagagagagcagagatacaagctcc 300
Qy 529 agccctcgcgctacaaccgcgcagatcgaggatataagatcacccggcgctgagcttac 588
Db 301 agccctcgcgctactaacgcgcagatcgaggatataagatcacccggcgctgagcttac 360
Qy 589 taactgtactgtcaggtgcaactttagatgagggagagagctgtctactcagagctgagctg 648
Db 361 taactgtactgtcaggtgcaactttagatgagggagagagctgtctactcagagctgagctg 420
Qy 649 ctggtgagatgtgtgtgctgacctgagctgcgctgcgagagagatctcaagcagctgcgcagt 708
Db 421 ctggtgagatgtgtgtgctgacctgagctgcgctgcgagagagatctcaagcagctgcgcagt 480
Qy 709 tcctctgggcccagctcgcgctcgcaggtgtctgcgctgtgtgcgctgcgagag 768
Db 481 tcctctgggcccagctcgcgctcgcaggtgtctgcgctgtgtgcgctgcgagag 540
Qy 769 tcctctctgagatcgcagacccctcccttggcccatctcaagagctgcgccccttcctcacc 828
Db 541 tcctctctgagatcgcagacccctcccttggcccatctcaagagctgcgccccttcctcacc 600
Qy 829 tacttcgagacttctcaggttctcagtgagggcctcgtgtctcccaagatgttccagagct 888
Db 601 tacttcgagacttctcaggttctcagtgagggcctcgtgtctcccaagatgttccagagct 660
Qy 889 gcgcgctccctcgcagacgtctctgagcccggtccctctgcgcccacccctcagcgct 948
Db 661 gcgcgctccctcgcagacgtctctgagcccggtccctctgcgcccacccctcagcgct 720
Qy 949 ctgttcacagacactgcgcccctccctctagagagctgcgctggcgctgttcaagtgatttcca 1008
Db 721 ctgttcacagacactgcgcccctccctctagagagctgcgctggcgctgttcaagtgatttcca 780
Qy 1009 tccacataaatacagatattcccaactctattctaacactcccccacgcgcacat 1063
Db 781 tccacataaatacagatattcccaactctattctaacactcccccacgcgcacat 835

RESULT 7
AAS03964 standard; DNA; 898 BP.
AAS03964;
26-SEP-2001 (first entry)
Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis; ocular neovascularisation; diabetic retinopathy; neovascular glaucoma; retinoblastoma; retinopathy of prematurity; retrolental fibroplasia; rubosis; uveitis; macular degeneration; arthritis; rheumatism; ds; corneal graft neovascularisation; psoriasis; metastatic condition; malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint; preneoplastic condition; myocardial angiogenesis; wound granulation; scleroderma; vascular adhesion; telangiectasia; ischaemia; human; atherosclerotic plaque neovascularisation; coronary atherosclerosis; peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR; fusion protein.
Homo sapiens.
Synthetic.
Key Location/Qualifiers
CDS 52..873
/*tag= a
/product= "Fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and human TWEAK extracellular domain"
WO200145730-A2.
28-JUN-2001.
19-DEC-2000; 2000WO-US34755.
20-DEC-1999; 99US-0172878.
10-MAY-2000; 2000US-0203347.
(IMMV) IMMUNEX CORP.
Wiley SR;
WPI: 2001-417975/44.
P-PSDB; AAU03499.
Modulating angiogenesis in a mammal for treating diseases mediated by angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or peripheral tissue, by administering antagonist or agonist of TWEAK receptor -
Example 1; Page 39-40; 46pp; English.
The sequence represents a DNA from the expression vector pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and the extracellular domain of human TWEAK. The fusion protein was used in the isolation of human TWEAK receptor (TWEAKR)-expressing clones from a COS cell human cDNA library. The TWEAK protein is a member of the tumour necrosis factor (TNF) family and induces angiogenesis. TWEAKR may therefore be used to screen for and develop TWEAKR agonists and antagonists for the modulation of angiogenesis, to be used in the treatment and diagnosis of human disease. The disorders mediated by angiogenesis include ocular disorders characterised by ocular neovascularisation such as diabetic retinopathy, neovascular glaucoma, retinoblastoma, retinopathy of prematurity, retrolental fibroplasia, rubosis, uveitis, macular degeneration and corneal graft neovascularisation, and inflammatory diseases such as arthritis, rheumatism and psoriasis. Other treatable diseases include malignant and metastatic conditions such as sarcomas and carcinomas, benign tumours and

CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;
Query Match 45.8%; Score 629.2; DB 22; Length 898;
Best Local Similarity 99.5%; Pred. No. 1e-124;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 232 agttgggagacgggacatgctgtccgcaagagctgcccagagagctgtgca 291
|||||
DB 250 agttgggagacgggacatgctgtccgcaagagctgcccagagagctgtgca 309
QY 292 gaggaggacagagaccgtcggaactgaatcccaagaaagaaagcagatctcg 351
|||||
DB 310 gaggaggacagagaccgtcggaactgaatcccaagaaagaaagcagatctcg 369
QY 352 ccttcctgaaccgaactagctgcctcgcaagtgcaactaaaggccgaaacacg 411
|||||
DB 370 ccttcctgaaccgaactagctgcctcgcaagtgcaactaaaggccgaaacacg 429
QY 412 gctcgaagagcagtcgacgccatcatgaagtcatcacaagactggaagagcag 471
|||||
DB 430 gctcgaagagcagtcgacgccatcatgaagtcatcacaagactggaagagcag 489
QY 472 caggcaggtgtggaagcagtgagtgctgaggagaaagcagaatacaagctcagc 531
|||||
DB 490 caggcaggtgtggaagcagtgagtgctgaggagaaagcagaatacaagctcagc 549
QY 532 ccttcgctacaacacgcagacatcgaggagttaagtaaccgagcttggtactaac 591
|||||
DB 550 ccttcgctacaacacgcagacatcgaggagttaagtaaccgagcttggtactaac 609
QY 592 ctgtactcgaagtgcaacttgatgaaggaaagctgttaccctgaagctgactctg 651
|||||
DB 610 ctgtactcgaagtgcaacttgatgaaggaaagctgttaccctgaagctgactctg 669
QY 652 gtggatggtgtgtgcccctcgctcgagaggaatttcagacacgagcagcttcc 711
|||||
DB 670 gtggatggtgtgtgcccctcgctcgagaggaatttcagacacgagcagcttcc 729
QY 712 ctcgagccccaagctccgctctcgccaggtgtctgggctgttgccctcgagcaggtcc 771
|||||
DB 730 ctcgagccccaagctccgctctcgccaggtgtctgggctgttgccctcgagcaggtcc 789
QY 772 tccctcgagatccgaacccctccctggggcccatctcaaggtgcccttctcaactac 831
|||||
DB 790 tccctcgagatccgaacccctccctggggcccatctcaaggtgcccttctcaactac 849
QY 832 ttcgagacttccaggttcaactcgaaggagccctgg 865
|||||
DB 850 ttcgagacttccaggttcaactcgaaggagccctgg 883
RESULT 8
ID AAV18599 standard; CDNA; 1168 BP.
AAV18599;
21-JUL-1998 (first entry)
Mus musculus tumour necrosis factor related ligand (TRELL) gene.
TRELL: tumour necrosis factor related ligand; tnfr; treatment;
KW cancer; autoimmune disease; immune system; stimulation; suppression;
CC graft rejection; ds.
OS Mus musculus.
XX

FH	Key	Location/Qualifiers
FT	CDS	2..679
FT		/tag=^
FT		a
FT		/note="tumour necrosis factor related ligand"
XX		
PN		WO9805783-A1.
PD		12-FEB-1998.
XX		
PF		07-AUG-1997; 97MO-US13945.
XX		
PR		18-MAR-1997; 97US-0040820.
PR		07-AUG-1996; 96US-0023541.
PR		18-OCT-1996; 96US-0028515.
XX		
PA	(BIOJ) BIOGEN INC.	
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.	
PI	Browning JL, Chicheportriche Y;	
XX		
PT	WPI: 1998-145619/13.	
XX	P-PSDB: AAM47524.	
XX		
XX	Tumour necrosis factor related ligand - useful for, e.g. treating cancer, auto-immune disease and immune responses to tissue grafts	
XX	Claim 2; Pages 45-46; 69pp; English.	
CC	The sequence is that encoding mouse tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon-gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. The DNA sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic defects, by introducing into cells, and expressing, therapeutically effective amounts of a vector, e.g. a virus comprising a gene encoding TRELL. It may also be of use in the preparation of prepare probes for screening natural/synthetic DNAs for TRELL-encoding sequences and for antisense therapy.	
CC		
CC	Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;	
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		</

D	b	242	cgagactatcgagcccatltagagttcattccctcggccgagacagagatgagacaaga	301
O	y	478	gggtgtgagccggagacgtgagttggcttggagagaaagccagaatcaacagctccagccctt	537
D	b	302	gggtgtgagttggagacgtgagttggcttggagagagaccaaatacaacagctccagccctt	361
O	y	538	cgctacaacccgccaagatccggggagattatagtcacccgggctcgggtctactactgttac	597
D	b	362	cgctacgcagcccgagatctggagattatcaacgcatcatcagggctcgggtctactactgttac	421
O	y	598	tgtcaggtgacacttgtatgttaggggaaagcgtgtactctgaactgaactgtgtgtgtgat	657
D	b	422	tgtcaggtgacacttgtatgttaggggaaagcgtgtactctgaactgaactgtgtgtgtgat	481
O	y	658	gggtgtgagccctcggccctcgtccggagaaattctcagccacgtcggccagattccctccg	717
D	b	482	gggtgtgagccctcggccctcgtccggagaaattctcagccacgtcggccagattccctccg	541
O	y	718	ccccagctcggctctgtccagagtgctctgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	777
D	b	542	ccccagctcggctctgtgtcagagtgctctgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	601
O	y	778	cgagttccggacacccctcccttggggccatctcaagcgtcggccctctctcaactcttgg	837
D	b	602	cgagttccggacacccctcccttggggccatctcaagcgtcggccctctctcaactcttgg	661
O	y	838	ctcttcacaggtctcaatgaggggcccgtgtctctccccaagctgtcccaaggtctcgggtcc	897
D	b	662	ctcttcacaggtctcaatgaggggcccgtgtctctccccaaggtctctcaactctccgtgtcc	721
O	y	898	-----ctcgcagactctctctggggaacccgggtccctctcgtcccaactcagcgtctt	952
D	b	722	agagagcatcacacaactccctccatcccaacccacatccctccacacccctc-gtgtctct	780
O	y	953	gtccagaacctggccctccctctctagagagctcgtcgtgtgtgtgtgtgtgtgtgtgtgt	1012
D	b	781	gtccagaacctggctctcc-----tcaagggcagccagagctgtgttcacatgttccatcc	837
O	y	1013	acataatacagtatctcccaactcttaactatacaactccccaacccgcccactctccactc	1072
D	b	838	-----acagacgtatctctgtcctctcttcaacatcccatcccaacaaatactccactc	891
O	y	1073	actagctccccaatccctgagaccccttggaggccccaagtatctcgactccccctgtggcca	1132
D	b	892	actagctccccaagccctac-----ttatccctgactcccccaccact	936
O	y	1133	cagagcccccagggatctgtgtctcaactgtactcgtgtggcaagagatggtgtccagaagacc	1192
D	b	937	caccgcgaccaaagtgttattgtactctgtgtac-----	968
O	y	1193	cacttcagagcaactaagggggctcggagacccgtcggcagagaagccaaagagactgggctag	1252
D	b	969	-----caggcaactgaagtgtggctcggagacccgtgtggcagagaagccaaagacactgggactag	1033
O	y	1253	gccaagagttccccaatgtgtgaggggcgagagaacaa--gacaagctcctccctgtgagaattc	1311
D	b	1024	gccaagagttccccaatgtgtgaggggcgagagaagcgtgggacaaagctcctccctcgtga---	1079
O	y	1312	ccgtgtgagtttttaaacagatatattttattattattattgtgtgacaaatgtgtga	1366
D	b	1080	ccgtgtgagtttttgaaa--agatactattttattattattattgtgtgacaaatgtgtga	1132
RESULT 9				
AAK23425				
ID AAK23425 standard; DNA: 701 BP.				
AC AAK23425:				
XX 18-JUN-1999 (first entry)				
XX Mouse TNRL3 DNA.				
XX				

KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KM developmental abnormality; gestational abnormality; prostate cancer;
 KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KM apoptosis; mouse; ss.

OS Mus sp.
 FT Key Location/Qualifiers
 FT CDS 1..636
 FT /tag= a
 FT /product= "TNRL3"

PN MO9911791-A2.

PD 11-MAR-1999.

PE 04-SEP-1998; 98WO-US18393.

PF 05-SEP-1997; 97US-0924634.

PI (UNIW) UNIV WASHINGTON.

PI Chaudhary PM;

DR WPI: 1999-205191/17.

DR P-PSDB; AAW93591.

PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities

PS Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Query Match 37.8%; Score 519.2; DB 20; Length 701;
 Best Local Similarity 87.3%; Pred. NO. 2.6e-101;
 Matches 569; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 220 ctggccggtgctcagtttgggagccgggcatcgctgcccagagagccagccagag 279
 DB 1 ctggccggtgctcagtttgggagccgggcatcgctgcccagagagccagccagag 60
 QY 280 gagctggtgag 339
 DB 61 gagctgag 120
 QY 340 cagagctcgctgcttccctgaacagagagagagagagagagagagagagagag 399
 DB 1 cagagctcgctgcttccctgaacagagagagagagagagagagagagagagag

DB 121 cagagctgtagctcttcttggaaacagtagtccggctcgaagaagtgtctctaagagc 180
 QY 400 cggaaacacgggtctgcaagaagagatcgcaagccattatagatctaccagacccgga 459
 DB 181 cggaaagcgcgctctcgagagagagagagagagagagagagagagagagagagag 240
 QY 460 caggacgagcgcag 519
 DB 241 cagagctgag 300
 QY 520 aacagctccagcctctcgagagagagagagagagagagagagagagagagagagag 579
 DB 301 aacagctccagcctctcgagagagagagagagagagagagagagagagagagagag 360
 QY 580 gggctctactactctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 639
 DB 361 gggctctactactctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 420
 QY 640 ctgagctgctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 699
 DB 421 ctgagctgctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 480
 QY 700 gggcagagcttccctcgagagagagagagagagagagagagagagagagagagag 759
 DB 481 gcaagcagcttccctcgagagagagagagagagagagagagagagagagagagag 540
 QY 760 cggcagagcttccctcgagagagagagagagagagagagagagagagagagagag 819
 DB 541 cggcagagcttccctcgagagagagagagagagagagagagagagagagagagag 600
 QY 820 ttccctactactctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 871
 DB 601 ttccctactactctgtagtcaagtagtcaagtagtcaagtagtcaagtagtcaagtag 652

RESULT 10
 AAT22190
 ID AAT22190 standard; CDNA to mRNA; 282 BP.
 XX
 AC AAT22190;
 XX
 DT 27-AUG-1996 (first entry)
 XX
 DE Human gene signature HUMG503761.
 XX
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09514772-A1.
 XX
 PD 01-JUN-1995.
 XX
 PF 11-NOV-1994; 94WO-JP01916.
 XX
 PR 12-NOV-1993; 93JP-0355504.
 XX
 PA (MATSU) MATSUBARA K.
 PA (OKUBO) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 1995-206931/27.
 XX
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 XX
 PS Claim 1; Page 1067; 2245pp; Japanese.
 XX

QY 506 aggaagccagatcacacagctccacccctctgcgtacacaccgacagatcgggagattca 565
DB 104792 VGGCCVGGGNNHNNNSGCCGCCVGGCGGNNHNNNSGCCGCCVGGCGGNNHNN 104733
QY 566 tagtcaccgggctgg 581
DB 104732 NSCGGCCCGCCVGGC 104717

RESULT 12
AA217263
ID AA217263 standard; cDNA: 1337 BP.
XX
AC AA217263;
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:4735.

Human; gene: gene expression product; diagnosis: therapy; probe:
detection; mapping: tissue typing; profiling: forensic; cancer:
genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

XX Homo sapiens.
XX OS
XX PN WO938972-A2.
XX PD 05-AUG-1999.
XX PF 28-JAN-1999; 99WO-US01619.
XX PR 03-APR-1998; 98US-0080666.
XX PR 28-JAN-1998; 98US-0072910.
XX PR 24-FEB-1998; 98US-0075954.
XX PR 31-MAR-1998; 98US-0080114.
XX PR 03-APR-1998; 98US-0080515.

XX PA (CHIR) CHIRON CORP.
XX PA (HXSE-) HXSEQ INC.

PI Cirvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX WPI: 1999-494092/41.

Novel human genes and their expression products which are
differentially expressed in different cell types

Claim 1; Page 2250-2251; 2479pp; English.

XX The present invention describes a library of human polynucleotides
XX comprising the sequences given in AA212532 to AA21779. Also described is
XX a method of detecting differentially expressed genes correlated with the
XX cancerous state of a mammalian cell, comprising detecting at least one
XX differentially expressed gene product in a test sample from a cell
XX suspected of being cancerous, where the gene product is encoded by one
XX of the 5248 polynucleotide sequences given in AA212532 to AA21779. The
XX polynucleotides can be used as a source of primers and probes, which can
XX be used for a variety of purpose, e.g. detection of expression levels,
XX mapping, tissue typing or profiling, forensics, genetic analysis and
XX detection of polymorphisms. Polypeptides encoded by the polynucleotides
XX can be used for raising antibodies for experimental, diagnostic and
XX therapeutic purposes. The polynucleotides may also be used to construct
XX arrays for diagnostics (which may be used to determine function of an
XX encoded protein); and to detect differences in expression levels between
XX two cells (e.g. to identify abnormal or diseased tissue in a human, to
XX identify a genetic predisposition or susceptibility to a disease such as
XX cancer). The polynucleotides of the invention are especially used in the
XX diagnosis, prognosis and management of colorectal cancer, breast cancer,

CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX
SQ Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

Query Match 4.4%; Score 61; DB 20; Length 1337;
Best Local Similarity 32.1%; Pred. No. 0.001;
Matches 145; Conservative 0; Mismatches 307; Indels 0; Gaps 0;

QY 690 ctcaagcaatcgagcgaatctccctcgtggcccaagctccgtctctgcagtggtcgtgct 749
DB 382 ccc 441
QY 750 gttgcccctgcggcagggtctctcctcgtgagtcgcacacccctccctgcgacatctca 809
DB 442 ccc 501
QY 810 ggtgcgccctctcctacactctcgaactctcgaagttcaactgaggggacctggtctc 869
DB 502 ncc 561
QY 870 cccacagtcgtccaggtcgtcggtccctcctcagacagctctctggcaccggtccctc 929
DB 562 ccc 621
QY 930 tgcacacccctcagcgtcttctgcacagactgcacctccctcctctagagctgctgg 989
DB 622 ccc 681
QY 990 cctgtcagtggttctcattccacataaaglatctccactctatcttacaactc 1049
DB 682 ncc 741
QY 1050 cccacagcgcacactctcactcactagctcctccatctccactgaccttgaggcccaag 1109
DB 742 ccc 801
QY 1110 tgatctgactcccccctgcgcacagacc 1141
DB 802 ncc 833

RESULT 13
AA202484/C
ID AA202484 standard; cDNA: 1000 BP.
XX
AC AA202484;

XX 19-MAY-2000 (first entry)

Human colon cancer cell line polynucleotide sequence SEQ ID NO:2475.

XX Human; colon cancer; tumour; diagnosis; gene expression product;
XX probe; detection; cancerous state; metastasis; identification;
XX breast cancer; oestrogen receptor-positive breast cancer; therapy;
XX oestrogen receptor-negative breast cancer; lung cancer; ss.

XX Homo sapiens.
XX OS
XX PN WO958675-A2.
XX PD 18-NOV-1999.
XX PF 13-MAY-1999; 99WO-US10602.
XX PR 14-MAY-1998; 98US-0085426.
XX PR 15-MAY-1998; 98US-0085537.
XX PR 15-MAY-1998; 98US-0085696.
XX PR 21-OCT-1998; 98US-0105234.
XX PR 27-OCT-1998; 98US-0105877.
XX PA (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.
XX Williams LR, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Ctkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LM, Stachne-Crain B;
XX
XX WPI; 2000-126369/11.
DR
XX
XX Polynucleotide library used to determine cancerous states of mammalian
PT cells -
PS
XX
XX Claim 1; Page 994; 1097pp; English.
CC
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
XX
SQ Sequence 1000 BP; 21 A; 53 C; 574 G; 22 T; 330 other;

Query Match 4.4%; Score 60.6; DB 21; Length 1000;
Best Local Similarity 36.1%; Pred. No. 0.0012;
Matches 171; Conservative 0; Mismatches 303; Indels 0; Gaps 0;

OY 668 cctgtgagctgctggaattctcagcagctgagcagctcctcgcgagccagctc 727
DB 911 cccggcc 852
OY 728 gctctgcagagtgctggtgtgtgctcgcgcagagtgctcctcgcagatcgcga 787
DB 851 ccccnnnnnnnccccccgggngccggccgngnncccccccccccgcttncgcccc 792
OY 788 cctcctcctgagccatctcagagctgcctcctcctcctcctcctcctcctcag 847
DB 791 ccccnnnnnnncc 732
OY 848 ttaactgaggggctgtgtctcccaagctgcagagctgcagagctgcagagc 907
DB 731 ccccccnnccgncnccnccnccnccnccnccnccnccnccnccnccnccnccncc 672
OY 908 tctctggagcagccgctcctcctcgcacacacacacacacacacacacacacac 967
DB 671 cgcgcnncnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccncc 612
OY 968 ctccctctagagctgctgctgctgctgctgctgctgctgctgctgctgctgct 1027
DB 611 cncnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccncc 552
OY 1028 tcccaactctatctaacactcccccagcagcagcagcagcagcagcagcagcagc 1087
DB 551 nccccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccncc 492
OY 1088 cctgagacctttagagcccccagtgatctgactcccccctgagcagagacccc 1141
DB 491 ccmgcnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccnccncc 438

RESULT 14

AA10594/c
ID AAA10594 standard; DNA; 10732 BP.
XX
XX
XX AAA10594;
AC
XX
XX 29-JUN-2000 (first entry)
DT
XX
XX Gene encoding a subunit of cellulose synthase.
DE
XX
XX Cellulose synthase; cellulose production; increase yield; ds.
KW
XX
XX Vigna angularis.
OS
XX
XX JP2000060568-A.
PN
XX
XX 29-FEB-2000.
PD
XX
XX 26-AUG-1998; 98JP-0239998.
PF
XX
XX 26-AUG-1998; 98JP-0239998.
PR
XX
XX (MIZU/) MIZUNO K.
PA (OJIP) OJI PAPER CO.
XX
XX WPI; 2000-342371/30.
DR P-PSDB; AAY85179.
XX
XX
XX A gene encoding a cellulose synthetic equipment - for the improvement
PT in the amount of cellulose synthesised in a plant body
XX
XX
XX Claim 2; Page 14-21; 32pp; Japanese.
XX
XX
XX This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesised by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant.
XX
XX
SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;

Query Match 4.3%; Score 59.2; DB 21; Length 10732;
Best Local Similarity 13.9%; Pred. No. 0.0039;
Matches 159; Conservative 454; Mismatches 515; Indels 13; Gaps 5;

OY 2 tgcatactgtagaacttgaatttcgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 61
DB 10503 TSTSRYYTYYTAAKTYTSRNCSTRNCSTRNCSTRNCSTRNCSTRNCSTRNCSTRNC 10444
OY 62 cccgggagtgagggggagtgaggcagagcagagcagagcagagcagagcagagcagagc 121
DB 10443 YSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRSTSRST 10384
OY 122 gccagagcgagagggggcgccgggggggggggggggggggggggggggggggggggg 181
DB 10383 TCCCYSDTBSRNSRST 10324
OY 182 tgggctcggcgccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 241
DB 10323 RCNSTYSCSRDYSNCCCYTSTYDASRCYSCNMCNSTRAKYSCYSTSTYDATTATY 10264
OY 242 gccggagatcgcgtctgcgcagagagcctccagagagagcagagagcagagagc 301
DB 10263 CYCTBCTC--SRNCTST 10206
OY 302 aggacccgtcgaactgaatccccaagagagagagagagagagagagagagagagagag 361
DB 10205 RGYSDAYDATBNCYYSYSTCCYTFSDCYTSRCTBSYDANCSTNCSTSTSTSTST 10146
OY 362 accgagactgctgctcgcgcagagagtgacactaaagcgggagagagagagagagag 421
DB 10145 ACSTYASTDNCSTRSTSTCTYAKYSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 10086

OY 422 cgaatcgagcccatlaagaattacacaagcactggaaggaaagcaaggaagt 481

Db 10085 C-TTNDNTRCIRASTYDASARCCSSSTYYRASRGIBSDCCBTBSRYCSYSTDAIRANCDTB 1002

OY 482 tgaacggagacagtgaabtgcttggaggaagccaagaalacacagctccagcccttcgcgt 541

Db 10026 CYTTBAKYPARCADAYAKRONNSTSRRAVSDSCYTCTRCNSHCKNSTYSYSTYPASTTB 9667

OY 542 acaacgccacgaalcggggagttatatagtaaccgggcttgggctctaaccgtactc 601

Db 9966 TTYCYCTBCSRKYSSRSYSTNCYSYCQYSRRSYSTTNNTGTCCTGYTSYNSTBYTS 9907

OY 602 agtgtcaccttgaatgaagggaagcgtctacctagaagctgaactcgtggygaigtg 661

Db 9906 YSSSTYS-----SRRYSRSGDSNGNCYNTNKNIDAMSTDTRYSCRCTYYISVSIDS 9855S

OY 662 ttgcttgcccctgcgctgccttggagaatltctcagcaactgcggcagcttccctcgggcc 721

Db 9854 TDSSTDSNSTTBSKCYYYTTBSTSDSTSRYCRSDYATBASBNSTNCCDARSRTBTB 9789S

OY 722 agctccgccccctgcagagtgtctcgggcttlttgccctcgtggccagggctccctgcgga 781

Db 9794 STNYCAACBYBDARCSRDTSVSSRGIYDANSTRYSRSSYSTYSDDSYSAHYCAKSTTB 9735S

OY 782 tcgcgaccccccttcgggccacataagcctgcccccttccataactt-ccgactc 840

Db 9734 TBCTDYADVADACYADVANCYSSDSTTBYBCSRGCCYYIAYSICSKRDARYACTYSYNS 9675S

OY 841 ttccaggttacatgaaggggcccttgctctcccaagctgcaccagctgcagagctcgccctcc 900

Db 9674 TCYDATBTTSRYSTTVSYSNICYDATTSRCTBYSTBTBTTSSRCAKTBDSTAKNSTST 9615S

OY 901 cgaaagctctcgtggcaccggcgc--ccctctgcgcccaacctagaocgctclttgctcag 959

Db 9614 YSTTRCTBYSRRSGYSYCSRRSRRCYSCYTOSDSTCVSTFTYYAVASCITTSRYASYDA 9555S

OY 960 acctgcgcccccttcgaagcgtcgtggcggccgtgtlcttaagtltttcaatccacataa 1019

Db 9554 SRSITSYSRCTTSTYSSTTDYDCDGITTTBNSTISSDSDBCTRYSDBRCRSNSTONC 9495S

OY 1020 tacagtatcccaactctatctatacaactcccccacggcccaactctcaactactagct 1079

Db 9494 YSCDSRSYSTTYDACYYDAKAICTYVSPDNCCNSTSRCYTNSTYSCSTRBSRNSTCCCT 9435S

OY 1080 ccccaatccctgaacctttaaggcccccaagtatctcgcactccccctcggccaagacc 1139

Db 9434 BTTSRGMCCTDAYDANSTSRVDYAPACYISTDASTBTYSCTBYSNTYDASSRYSC 937S

AA02488 :
1140 c 1140
9374 y 9374

RESULT 15
AAA02488/c ID AAA02488 standard; cdna; 1218 bp.

XX AAA02488:
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2479.
DT 19-MAY-2000 (first entry)
XX XX
KW Human; colon cancer; tumour; diagnosis; gene expression product;
PM Probe; detection; cancerous state; metastasis; identification;
KM breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX OS Homo sapiens.
XX XN
XN W09958675-A2.

	PX	18-NOV-1999.	
	PF	13-MAY-1999;	99NM-USt0602.
	XX		
	PR	14-MAY-1998:	98US-0085426.
	XX		
	PR	15-MAY-1998:	98US-0085537.
	XX		
	PR	15-MAY-1998:	98US-0085696.
	XX		
	PR	21-OCT-1998:	98US-0105234.
	XX		
	PR	27-OCT-1998:	98US-0105877.
	PA	(CHIR) CHIRON CORP.	
	PI	(HSE-) HYSEQ INC.	
	PI	Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,	
	PI	Reinhard C, Glase K, Randazzo F, Kennedy GC, Pot D, Kassam A;	
	PI	Lemson G, Dimanac R, Crkvenjakov R, Dickson M, Dimanac S, Labat I,	
	PI	Leshkowitz D, Kita D, Garcia V, Jones LM, Stache-Craln B;	
	DR	WPI; 2000-126369/11.	
	PT	Polynucleotide library used to determine cancerous states of mammalian	
	PT	cells -	
	XN		
	PS	Claim 1; Page 995-996; 1097pp; English.	
	CC	AAA0010 to AAA02716 represent polynucleotides isolated from cDNA	
	CC	libraries constructed from human colon cancer cell lines. The present	
	CC	invention also describes a method of detecting differentially expressed	
	CC	genes correlated with a cancerous state of a mammalian cell, comprising	
	CC	detecting at least one differentially expressed gene product in a test	
	CC	sample derived from a cell suspected of being cancerous, where detection	
	CC	of the differentially expressed gene product is correlated with a	
	CC	cancerous state of the cell from which the test sample was derived.	
	CC	The polynucleotide sequences can be used in a method for detecting	
	CC	differentially expressed genes correlated with a cancerous state of a	
	CC	mammalian cell. The polynucleotides can also be used as probes for	
	CC	detecting and mapping related genes. They can be used in diagnosis and	
	CC	prognosis of diseases and disorders (e.g., identification of	
	CC	pre-metastatic or metastatic cancerous states, stages of cancer, or	
	CC	responsiveness of cancer to therapy). This is particularly for breast	
	CC	cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-	
	CC	negative breast cancer, lung cancer, and colon cancer.	
	SQ	Sequence 1218 BP; 9 A; 31 C; 494 G; 37 T; 647 other;	
	Query Match	4.2%; Score 58; DB 21; Length 1218;	
	Best Local Similarity	27.1% Pred. No. 0.0044;	
	Matches 118; Conservative %	0; Mismatches 318; Indels 0; Gaps 0;	
OY	DB	704 ccagttccctcgtgggcccaggctcgacctgcgcgaagtgtctcgtagtcttgaccctggcgc	1048
OY	DB	CCCNNNNNNNNCCCCCNCNNNNNNNNCNCCNNNNNNNNNNNNNNNNCCGCCNNCCNNCN	1048
OY	DB	caagggttcctccctcgaga tcgcacaactccccctgggcccatctcaaagctgcgacctcc	823
OY	DB	CNNNNNNNNNNNNNNNNNNNNNNCCCCCCCCCCCGCCNNNNNNNNCCGNNNNNNCCCCCCCC	988
OY	DB	tcaacttacitcgaactcttcaggltcacgtagagggcccttgtctccccaagtgcctcc	883
OY	DB	NNCNC(CCCNNNNNNNNNNNNNC(NC(NC(NNNNNNNNNCCCCCCCCCNCNNCCGCNC(CNNNNNNNN	928
OY	DB	aggtgcgcggtcctccctcgacagctcctcttgggaaacccggtctccctctgccaacctta	943
OY	DB	NNCCCCCCCCCCCCCCCCCNCNNNNNNNNCCGNCNNCC(CNNNNNNNNNNNNNNNNCCGNNN	868
OY	DB	ccgctcttgttcagagactgcctccctctaagaaggtcgtagctgtgtcaagtgtc	1003
OY	DB	N(CNNNNNNNNNNNNNNNNNNNNNNCCCCCCCCCCGN(CNNNNCC(CCCCC(GGGCCCC(CCNN	808
OY	DB	ttcataccataaatcaglatlcccatcttatctlaaacactcccaacgccaact	1063
OY	DB	TTCNNNNNNNNCCCCCNCNNNNNNNNCC(NC(NC(NC(NC(NNNNNNNNNNNNNNNNNCCGNNNNNN	748

THIS PAGE BLANK (USPTO)

102 CAGAGGAGGACCAGGACCCGTCGGAACCTGAATCCCCAGACAGAGAAAGGC 15

114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAl 130
|||||
152 CAGATGCTCGCGCTTCTCTGAACCGACTAGTGGCCCTCGCAGAAGTCC 201
|||||
130 aProlySGlyArglyThrArgAlaArgAlaIleAlaIleHisTyrG 147
|||||
202 ACCTAAAGCCGAAACACGGGCTCGAAGACCGATCGCAGCCCATATG 251
|||||
147 LValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
|||||
252 AGGTCATCAGCAGCCTGGACAGACGAGCCAGCAGGTGTGGACGG 301
|||||
164 ThrValSerGlyTyrPgluGlnAlaArgIleAsnSerSerProLeuAr 180
|||||
302 ACAGTAGAGTGCTGGGAGAACCCAGAAATCAACAGCTCCAGCCCTCTGCG 351
|||||
180 gTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrT 197
|||||
352 CTACAACCCGACAGATCGGGAGTTTATAGTCAACCCGGGCTGGGCTCTACT 401
|||||
197 yLeuTyrCysGlnValHisPheAspGlyGlyValAlaValTyrLeuLys 213
|||||
402 ACCTGACTGTCAGGTGCACCTTTGATGAGGAGGAGGCTCTTACCTGAAG 451
|||||
214 LeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluP 230
|||||
452 CTGAGACTTCTGTGGATGTGTGCTGGCCCTGCGCTCGCTGAGAGAATT 501
|||||
230 eSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgCysGlyVal 247
|||||
502 CTCAGGCACTGGCGGCAGTTCCTCGGGCCCAAGCTCCGCTGTGCAGG 551
|||||
247 aLserGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgTyr 263
|||||
552 TGTCTGGGCTGTGGCCCTGGCCGACAGTTC.TCCCTGGGAGATCCGACAC 600
|||||
264 LeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLe 280
|||||
601 CTCCTCGGCGCCCATCTCAAGGCTGGCCCTTCCTCCTACTTCGAGCT 650
|||||
280 uPheGlnValHis 284
|||||
651 CTTCACAGGTTTCAC 663
|||||
seq_name: gb_est2:BI870393
seq_documentation_block:
ACCS BI870393 728 bp mRNA linear EST 11-OCT-2001
DEFINITION 603395641F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',
mRNA sequence.
ACCESSION BI870393
VERSION BI870393.1 GI:16044066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 728)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@db-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034 row: a column: 12
High quality sequence stop: 728.
Location/Qualifiers

source 1. 728
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405459"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIH_MGC Library."
BASE COUNT 125 a 240 c 227 g 136 t
ORIGIN
alignment_scores:
Quality: 1060.50 Length: 221
Ratio: 4.842 Gaps: 1
Percent Similarity: 99.095 Percent Identity: 99.095
alignment_block:
US-09-245-198a-4 x BI870393 ..
Align seg 1/1 to: BI870393 from: 1 to: 728
65 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValaSerLeuGlySe 81
|||||
2 CTGGCCCTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 51
|||||
81 rArgAlaSerLeuSerAla...GlnGluProAlaGlnGluGluLeuVala 97
|||||
52 CCGGGCATGCTGTGCGCCAGCAGAGAGGCTGCCAGAGAGAGTGTGG 101
|||||
97 lAGluGlnAspGlnAspProSerGlyLeuAsnProGlnThrGluGluSer 113
|||||
102 CAGAGAGAGCAGAGACCCGCTGGAATCTGAATCCCAAGCAAGAAAC 151
|||||
114 GlnAspProAlaProPheLeuAsnArgLeuValaArgProArgSerAl 130
|||||
152 CAGATGCTCGCGCTTCTCTGAACCGACTAGTGGCCCTCGCAGAAGTCC 201
|||||
130 aProlySGlyArglyThrArgAlaArgAlaIleAlaIleHisTyrG 147
|||||
202 ACCTAAAGCCGAAACACGGGCTCGAAGACCGATCCGACCCCATATG 251
|||||
147 LValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
|||||
252 AGGTCATCAGCAGCCTGGACAGACGAGCCAGCAGGTGTGGACGG 301
|||||
164 ThrValSerGlyTyrPgluGlnAlaArgIleAsnSerSerProLeuAr 180
|||||
302 ACAGTAGAGTGCTGGGAGAACCCAGAAATCAACAGCTCCAGCCCTCTGCG 351
|||||
180 gTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrT 197
|||||
352 CTACAACCCGACAGATCGGGAGTTTATAGTCAACCCGGGCTGGGCTCTACT 401
|||||
197 yLeuTyrCysGlnValHisPheAspGlyGlyValAlaValTyrLeuLys 213
|||||
402 ACCTGACTGTCAGGTGCACCTTTGATGAGGAGGAGGCTCTTACCTGAAG 451
|||||
214 LeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluP 230
|||||
452 CTGAGACTTCTGTGGATGTGTGCTGGCCCTGCGCTCGCTGAGAGAATT 501
|||||
230 eSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgCysGlyVal 247
|||||
502 CTCAGGCACTGGCGGCAGTTCCTCGGGCCCAAGCTCCGCTGTGCAGG 551
|||||
247 aLserGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgTyr 263
|||||
552 TGTCTGGGCTGTGGCCCTGGCCGACAGTTC.TCCCTGGGAGATCCGACAC 600
|||||

264 LeuProTPrAlaHisLeuLysAlaAlaProPheLeuThrTyPheGlyLe 280
|||||
601 CTCCTGTGGGCGCCATCTCAAGCTGC.CCCTTCTCATCTACTTCGACT 649
280 uPheGlnValHis 284
|||||
650 CTCACAGTTTCAC 662
seq_name: gb_est2:BI819200
seq_documentation_block: 777 bp mRNA linear EST 04-OCT-2001
LOCUS BI819200 603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',
DEFINITION mRNA sequence.
ACCESSION BI819200
VERSION BI819200.1 GI:15930750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 777)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1437 row: 1 column: 03
High quality sequence stop: 759.
Location/Qualifiers
1. 777
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (InVivoGen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 140 a 268 c 215 g 154 t
ORIGIN
alignment_scores: 979.00 Length: 188
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-245-198a-4 x BI819200 ..
Align seg 1/1 to: BI819200 from: 1 to: 777
97 AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGln 113
|||||
14 GCAGAGGAGGACGACGCCCTCGAATCGAATCCCGACGAGAAAGAAAG 63
|||||
113 rGlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerA 130

|||||
64 CCAGGATCTCGCGCTTCTGTAACCGCACTGATGGCGCTCGCAAGAACTG 113
130 laProLysGlyArgLysThrArgAlaArgArgAlaIleAlaAlaHisTyr 146
|||||
114 CACCTAAAGGCCGGAACACAGGCGCTCGAAGAGGATCGACGCCCATAT 163
147 GluValHisProArgProGlyGlnAspGlyValGlnAlaGlnAlaValAspG 163
164 GAAGTTCTATCCACGACCTCGACAGGACGAGCGGACGAGGTGTGGACGG 213
163 yThValSerGlyTyrPglGlnGlnAlaArgIleAsnSerSerProLeuA 180
214 GACAGTGAAGTGCTGGGAGGAAGCAGAAATCAACAGCTCCAGCCCTTCG 263
180 rGlyTyrAsnArgGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyr 196
264 GCTACAAACCGCCAGATCGGGAGATTATAGTCACCCGGGCTGGCTTAC 313
197 TyrLeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLy 213
314 TACCTGTACTGTCAAGTGCACATTGTATGAGGGAAGGCTGTCTACTTGA 363
213 sleuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGlnUp 230
364 GCTGAGCTTGTCTGGATGGATGTGTGCTGGCCCTCGCTGAGGAGAT 413
230 heserAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgCysGln 246
414 TCTCAGCCACATCGGCGGCTTCCCTCGGCGCCAGCTCCGCTTGCACG 463
247 ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgTh 263
464 GTGCTGTGGGCTGTGGCCCTCGGCGGAGGTCTTCCCTCGGAGATCCGAC 513
263 rLeuProTPrAlaHisLeuLysAlaAlaProPheLeuThrTyPheGlyL 280
514 CTCGCCCTGGGCGCCATCTCAAGGCTGCCCTTCTCTCACTACTTGGAC 563
280 eupheGlnValHis 284
|||||
564 TCTTCCAGGTTTCAC 577
seq_name: gb_est2:BI766766
seq_documentation_block: 834 bp mRNA linear EST 25-SEP-2001
LOCUS BI766766 603056866F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
DEFINITION mRNA sequence.
ACCESSION BI766766
VERSION BI766766.1 GI:15758344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 834)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1517 row: c column: 18
High quality sequence stop: 772.
Location/Qualifiers
1. 834
FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_1db="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NOTI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb. Insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT      154 a      287 c      223 g      170 t
ORIGIN

```

```

Alignment_scores:
Quality: 962.00      Length: 194
Ratio: 5.010      Gaps: 1
Percent Similarity: 98.969      Percent Identity: 98.969

```

alignment_block:
US-09-245-198a-4 x B1766766 ..

Align seg 1/1 to: B1766766 from: 1 to: 834

```

92 GINGLUGLULeuValAlaGLUGLUSpGlnAspProSerGluLeuAsnPr 108
|||||
3 CAGGAGAGACTGGTGGCAGAGAGAGACCGAGAGCCCTCGAAGTGAATCC 52
|||||
108 oGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValA 125
|||||
53 CCAGACAGAGAAAGCCAGAGATCTCGCCCTTCTGAAACCGACTGATTC 102
|||||
125 rGProArGArSerAlaProLysGlyArgLysThrArgAlaArgArgAla 141
|||||
103 GGCCTCGCAGAGTGCACCTAAAGCCGGGAAACAGGGGCTCGAAGAGCG 152
|||||
142 TLeAlaAlaHIStyrGluValAlaHISProArGPro.GlyGlnAspGlyAlaG 158
|||||
153 ATCCGAGCCCATTTATGAACTTCAACACCGCTGAGAGAGAGAGCGCG 202
|||||
158 TAlaGlyValAlaSpGlyThrValSerGlyTTPGluGluAlaArgIleAsn 174
|||||
203 AGGCAAGTGTGACGGGACAGTGAAGTGGCTGGGAGAGACCAATCAAC 252
|||||
175 SerSerSerProLeuArGtyrAsnArGlnIleGlyGluPheIleValTh 191
|||||
253 AGCTCAGCCCTCGCGTACAAACCGCAGATCGGGAGCTTATGACTCAC 302
|||||
191 rArgAlaGlyLeuTyrTyrLeuTyrCysGlnValAlaHISpHeAspGlyL 208
|||||
303 CCGGCGCTGGGCTTACTACTGATGTCAGGTGCACTTGTGATGAGAGGA 352
|||||
208 yAlaValAlaTyrLeuLysLeuAspLeuLeuValAspGlyValLeuAlaLeu 224
|||||
353 AGGCTTCTACTGAAAGCTGGACTTCTGGTGGATGTGTGCTGGCCCTG 402
|||||
225 ArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProG 241
|||||
403 CGCTGCTGAGAGAAATTCACGCACTGGCGCCAGTTCCCTGGGGCCCA 452
|||||
241 nLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerS 258
|||||
453 GCTCCGCTCTGCGCAGGTGTGGGCTGTGGCCCTGCGGCCAGGTGC.T 501
|||||
258 rLeuLeuArgTyrLeuArgThrLeuProTyrAlaHISLeuLysAlaAlaProPhe 274
|||||
502 CCCTGGGATCCGACACCTCCCTGGGGCCCATCTCAAGGCTGCCCTTC 551

```

```

275 LeuThrTyrPheGlyLeuPheGlnValHIS 284
|||||
552 CTCACCTACTCGAGACTCTTCCAGGTTTAC 581

```

seq_name: gb_est2:B1596681

seq_documentation_block:

```

LOCUS      B1596681      828 bp      mRNA      linear      EST 07-SEP-2001
DEFINITION 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
            mRNA sequence.
ACCESSION  B1596681
VERSION    B1596681.1 GI:15489620
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

1 (bases 1 to 828)
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Mikhail Palkovits, M.D., Ph.D.
            CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: L16M11722 row: k column: 13
            High quality sequence stop: 776.

```

FEATURES

source

1..828

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/clone_1db="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptPR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TGTGTTTGTGTTTGTGTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

```

BASE COUNT 155 a 278 c 223 g 172 t

alignment_scores:

```

Quality: 934.00      Length: 180
Ratio: 5.189      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.444

```

alignment_block:
US-09-245-198a-4 x B1596681 ..

Align seg 1/1 to: B1596681 from: 1 to: 828

```

105 GULeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAs 121
|||||
44 GAACCTGAATCCCGACAGCAAGAAAGCCAGATCCGCGCTTCTCTGA 93
|||||
121 nArgLeuValArgProArgSerAlaProLysGlyArgLysThrArgA 138
|||||
94 CCGACTAGTTGGCCCTCGCAGAAAGTGCACCTAAAGCCGGAAACAGCGG 143
|||||
138 TArGArGAlaIleAlaHIStyrGluValAlaHISProArGProGlyGln 154

```

144 CTCGAGAGCGATCGACGCCATTATGAAATTTCACCGACCTGCAGCAG 193
155 AspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTTPGluGluAl 171
194 GACGAGCGCGGAGGAGTGAGTGGACGATGAGTGGCTGGAGGAGC 243
171 AArgIleAsnSerSerProLeuArgTyrAsnArgGlnIleGlyLup 188
244 CAGATCAACAGCTCCAGCCCTTCGCTACACCGCCAGATCGGGGAGT 293
188 heIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 204
294 TTATAGTACCCGGGCTGGCTACTACTACTACTACTACTACTACTACTT 343
205 AspGluGlyValAlaValTyrLeuLeuLeuLeuLeuValAspGlyVal 221
344 GATAGGGGAGGAGCTGTCTACCTGAACTGAGCTTGGTGGATGTGTGT 393
221 IleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAspSerL 238
394 GCTGGCCCTGGCTGCTCGGAGGAATTCACGCCACCTGGCCAGTTCC 443
238 euGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArg 254
444 TCGGGCCCGACGCTCGGCTGTCCAGGAGTGTGGGCTGGCCCTGGCG 493
255 ProGlySerSerLeuArgIleArgThrLeuProThrPalahisLeuVal 271
494 CCAGGGTCTCTCCGCGGATCCGACCCCTCCGCGGCCATCTCAAGCC 543
271 aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
544 TGCCCCCTCTCTCACTACTTGGACCTTCCAGGTTTAC 583

seq_name: gb_est2:BF577781

seq_documentation_block: 918 bp mRNA linear EST 12-DEC-2000
LOCUS BF577781
DEFINITION 602092080F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206595 5',
mRNA sequence.

ACCESSION BF577781 GI:11651493
VERSION BF577781
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9767 row: P column: 20
High quality sequence start: 17
High quality sequence stop: 724.
Location/Qualifiers

FEATURES
source 1..918
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4206595"
/clone_1ib="NCI_CGAP_Co24"
/lab_host="DH10B (r1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library.
BASE COUNT 153 a 292 c 268 g 204 t 1 others
ORIGIN

alignment_scores:
Quality: 913.00 Length: 218
Ratio: 4.411 Gaps: 0
Percent Similarity: 94.954 Percent Identity: 88.073

alignment_block:
us-09-245-198a-4 x BF577781 ..

Align seg 1/1 to: BF577781 from: 1 to: 918

67 LeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAl 83
2 CTGGCCCTGGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51
83 aSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluA 100
52 AACCTGTCTGCCAGGAGCCTTCTCAGAGAGACTGACACAGACAGACACC 101
100 spGlnAspProSerGlnLeuAsnProGlnGluGluGluGluGluGluGlu 116
102 CGCGGAGCCCGCTGAACTGAAATCCAGACAGAGAGAGAGAGAGAGTGG 151
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaPolySGL 133
152 GTACCTTCTTGAACAACACTGTCGGGCTCGAAGAGTGTCTTAAGG 201
133 YArgLysThrArgAlaArgArgAlaIleAlaAlaHisTyrGlnValHisP 150
202 CCGGAGGCGCGGCT.CCGCGAGCTATTGACGCCATTATGAGGTGATC 250
150 roArgProGlnGlnAspGlyAlaGlnAlaGlnAlaValAlaAspGlyThrValSer 166
251 CTGGCCCGAGCAGAGATGAGACACAGAGAGTGTGAGTGGAGCAGTGTGT 300
167 GlnTyrGluGluAlaArgIleAsnSerSerProLeuArgTyrAsnAr 183
301 GCTGGGAGAGAGACCAAAATCAACAGCTCCAGCCCTGTGGCTACGACCG 350
183 gGlnIleGlyLupheIleValThrArgAlaGlyLeuTyrTyrLeuTyrC 200
351 CCAGATTGGGGAATTTACATCAGCAGGCGTGGGCTCTACTACTACTACT 400
200 YSGLNValAlaHisPheAspGluGlyLysAlaValTyrLeuLeuLeuAspLeu 216
401 GTCAAGTGCCTTGTGATGAGGAAAGGCTGTACCTGAACTGAGACTGTG 450
217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaTh 233
451 CTGTGGAAGGAGTGTGCTGGGCTCGGCTGCTGGAAGAATTCACAGCAC 500
233 rAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
501 AGCAGCAAGCTCTCTGGGCCCCAGCTCCGTTTGTGCAGAGTGTCTGGGC 550
250 euLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTyr 266
551 TGTTCGCGCTGGCGCAGGTCT.TCCCTTGGATCCGACCCCTCCCTGG 599
267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
600 GCTCATCTTAAGGCTGC.CCCTTCTAATACTTTGGACTTTCAAGT 648
283 His 284
649 TCAC 652

seq_name: gb_est1:AW763237

seq_documentation_block:

LOCUS AW763237 561 bp mRNA linear EST 04-MAY-2000

DEFINITION ut70d09.y1 NCI-CGAP Mam3 Mus musculus cDNA clone IMAGE:155633.5' similar to TR:054907.054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS ; mRNA sequence.

ACCESSION AW763237

VERSION AW763237.1 GI:7695174

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 561)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Other-ESTs: ut70d09.x1

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/lresources.shtml

MG1:1058389

Seq primer: -40RP from Gldco

High quality sequence stop: 433.

FEATURES

source

1..561

/organism="Mus musculus"

/strain="129 - C57/Bl6 - EVBN"

/db_xref="taxon:10090"

/clone="IMAGE:315633"

/clone_lib="NCI CGAP_Mam3"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies, Investigators providing samples: lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others

ORIGIN

Alignment_scores:

Quality: 825.00 Length: 187

Ratio: 4.609 Gaps: 0

Percent Similarity: 95.722 Percent Identity: 86.096

Alignment block:

US-09-245-198a-4 x AW763237 ..

Align seg 1/1 to: AW763237 from: 1 to: 561

41 SerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeuLeu 57

|||||

1 AGCCAGAGCCGGAGGGGGCCGGGGGAGCCGGGACCCCTGCTGGC 50

57 lProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLeuGlyLeuLeu 74

|||||

51 CCGCGTGGTGAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 100

74 euAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluPro 90

|||||

101 TGGTCGTGCTCAGCTGGGAGCTGGGCAAGCTGCTGCCAGAGCCT 150

91 AlAGInGluGluLeuValAlAGInGluAspGlnAspProSerGluLeuAs 107

|||||

151 TCTCAGAGAGAGCTCAGCAGCAGAGACCGCGGAGCCCTCGAACTGA 200

107 nProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 124

|||||

201 TCCCGACAGAGAGAAAGCCAGATGTGTACTCTTCTTGGAACTAG 250

124 AlArgProArgSerSerAlaProGlyArgGlyThrArgAlaArgArg 140

|||||

251 TCCGGCTCTGAGAGAGTCTCTTAAAGCCGGAGGCGGCTGCGCA 300

141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAl 157

|||||

301 GCTATTGTCAGCCCATTTAGAGTTTATCTCTGCGCCAGACAGATGGAGC 350

157 agInAlaGlyValAspGlyThrValSerGlyTyrPheGluAlaArgIleA 174

|||||

351 ACAAGAGAGTGTGATGGAGACAGTGAGTGGGAGAGACCAAAATCA 400

174 snSerSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheLeuVal 190

|||||

401 ACAGTCCAGCCCTCTGCGCTACGACCCGACATGGGAAATTTACAGATC 450

191 ThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlu 207

|||||

451 ATCAGAGGCTGGGCTACTACTACTGTCTGAGTGCACATTTGATGAGAG 500

207 yLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLe 224

|||||

501 AAAGCTGTCTACTGAAAGCTGAGCTGCTGCTGTAAGCTGTGGTGGCC 550

224 euArgCysLeu 227

|||||

551 TCGCGTGGCTG 561

seq_name: gb_est2:BT62908

seq_documentation_block:

LOCUS BT62908 785 bp mRNA linear EST 25-SEP-2001

DEFINITION 60304/966F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188168.5', mRNA sequence.

ACCESSION BT62908

VERSION BT62908.1 GI:15754486

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 785)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

plate: L14M11470 row: c column: 17

High quality sequence stop: 647.

FEATURES

source

1..785

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5188168"

/clone_lib="NIH_MGC_116"

/lab_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."

BASE COUNT 124 a 228 c 274 g 159 t
ORIGIN

alignment_scores:

Quality: 742.50 Length: 258
Ratio: 3.519 Gaps: 18
Percent Similarity: 81.783 Percent Identity: 75.581

Alignment block:
US-09-245-198a-4 x B1762908 ..

Align seg 1/1 to: B1762908 from: 1 to: 785

```

25 G1yAlaValArgGlnAlaGlnProAlaProMeta1aAlaArgArgS 41
|||||
29 GGGCGGTGAGTCACTGCACAGCCCCCGCCATGCTCTCCGCGA 78
41 erGlnArgArgArgGlyArgArg.Gly.GlnProGly...Thra1aLeu1 56
|||||
79 GCCAGAGCGGAGGGGCGCGGTGAGACCTGGCCACTGCCCTCT 128
56 euValaProLeuAlaLeuGly...LeuGlyLeuAla.LeuAlaCysLeu.. 70
|||||
129 TGGTCCCTGCTGCTGCTGCTGCTGGCGCTGCTGCTGCTGCT 178
71 .....GlyLeuLeuAlaValSerLeuGlySerArgAlas 84
|||||
179 GCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 222
84 erLeuSerAlaGlnGlnProAlaGln.GlnGlnLeuValAlaGlu...G1 99
|||||
223 CTGTCGCGCCAGTGAAGCTGCCAGTGAAGCTGGTTGGCAGAGTGA 272
99 uASGlnAspProSer...GluLeuAsnProGlnThrgLnuSerGlnA 115
|||||
273 TGACCACTGACCCCTGCTGGAACGAAATCCCAACAGAAAGCCAG 322
115 sPProAlaProPheLeuAsnArg..LeuValArgProArgArgSerAlaPr 131
|||||
323 ATCTGCGCCCTTCTCTGACCTGACTACTTGGCTCGCCACAAGTGCAC 372
131 OlySGlyArgGlyThrArgAlaArgArgAla1eAla1ah1stYrgLuv 148
|||||
373 TAAAGCGCGGAACACGCGCTCGAAGAGCATGCAACCATATATAG 422
148 a1h1sPProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThr 164
|||||
423 TTTCATCCACGACCTGACAGAGCAGCGAGCGAGGTGAGCGGAC 472
165 Val.SerGlyTyrPgluGlnAlaArgIleAsnSerSerSerProLeuArgT 181
|||||
473 GTTAGTGCGCGGAGAGAGCCAGATACACACTCCAGCCCTTGGCGCT 522
181 yrasnArgGlnIleGlyGlnPheIleValThraArgAlaGlyLeuTyrTyr 197
|||||
523 ACAACCGCAGATCGGGAGTTTATAGTACCCGCGCTGAGTAC 572
198 LeuTyrCysGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLe 214
|||||
573 CTGACTGTCTAGGTGACACTTTGATGAGGGAAGGCTGTACTGAAACT 622

```

```

214 uASpLeuValAspGlyValLeuAla...LeuArgCysLeuGlnGluP 230
|||||
623 GGACTTGTGTTGGATGTGTGCTGGCGCTGTGCTTCTTGGAGACAT 672
230 heSerAlaThra1aAla.SerSerLeu.GlyProGlnLeuArgLeuCysG 246
|||||
673 TCTCAGGACACTCGCGGAGTTCCTCGGGGCCAGCTCCGCTGCTGCC 722
246 lnyValSer.GlyLeuLeu.AlalaArgProGlySerSerLeuArgIleA 262
|||||
723 AGGTGTCTGGGGCTGTGGCTTGGGGGAGGTCCTGCTGATGCC 772
262 rg 262
||
773 GG 774

```

seq_name: gb_est2:BG686319

seq_documentation_block:

LOCUS BG686319 587 bp mRNA linear EST 01-MAY-2001
DEFINITION 602638232F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766071 5',

ACCESSION BG686319

VERSION BG686319.1 GI:13917716

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 587)

NIH-MGC http://imgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

tissue procurement: Louis M. Staudt, M.D., Ph.D.

cDNA library preparation: Ling Hong/Rubin Laboratory

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM1625 row: P column: 08

High quality sequence stop: 587.

Location/Qualifiers

1..587

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4766071"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site.1: XhoI;

Site.2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCAAGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the Laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

BASE COUNT 108 a 187 c 205 g 87 t

ORIGIN

alignment_scores: Quality: 738.00 Length: 156
Ratio: 4.855 Gaps: 2
Percent Similarity: 97.436 Percent Identity: 97.436

alignment_block: US-09-245-198a-4 x BG686319 ..

gene 1. .433
/gene="Tnfsf12"
CDS <1. .433
/gene="Tnfsf12"
/note="data source:MCD, source key:MGI:1196259,
evidence:ISS
putative
tumor necrosis factor (ligand) superfamily, member 12"
/codon_start=2
/protein_id="BA32249.1"
/db_xref="GI:12861641"
/translation="IAAHYEVHRRPQDAQAQAVDCTVSGMEETKINSSPLRYDRI
GEFTVIRAGLYLYCVHFEDEKAVYIKLIDLNVGVIALCEEFSAATASSPQDLR
LCOVSGLLPRPSSLRIRITLPAHLKAAPFLTYFGILFOVH"
BASE COUNT 228 a 300 c 248 g 256 t 1 others
ORIGIN

alignment_scores:
Quality: 707.00 Length: 143
Ratio: 5.086 Gaps: 0
Percent Similarity: 97.203 Percent Identity: 94.406

alignment_block:
US-09-245-198a-4 x AK020909 ..

Align seg 1/1 to: AK020909 from: 1 to: 1033

142 IleaAlaAlaHsTyrGluValHisProArgProGlyGlnAspGlyAlaG1 158
|||||
2 ATTCAGAGCCATTTAGAGTTTCATCTCGCCAGACAGATGAGACACA 51
|||||
158 nAlaGlyValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsn 175
|||||
52 AGCAGGTGTGATGGACAGTGAAGTGGTGGAGACCAAAATCAACA 101
|||||
175 erSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThr 191
|||||
102 GCTCCAGCCCTCTGGCTACAGCCGACGATTTGGGATTTACAGTATC 151
|||||
192 ArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyLy 208
|||||
152 AGGGCTGGGCTCTACTACTGTCTAGCTGACGCTGACTTTGATGAGGAAA 201
|||||
208 sAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuA 225
|||||
202 GGCTGTCTACTGAAGCTGAGCTGTGTGAGACGTTGCTGGCCCTGC 251
|||||
225 rgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGln 241
|||||
252 GCTGCTGTGAGAAATTCACAGCCACAGCAGACGCTCTCCGGGCCAG 301
|||||
242 LeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerS 258
|||||
302 CTCGGTTGTGCTCAGAGTGTCTGGGCTGTGGCGCTGCGCCAGGGCTTC 351
|||||
258 rLeuArgLeuArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheL 275
|||||
352 CTTTGAGTCCGACACCTCCCTGGGCTCATTTAAAGCTGCCCTCCCTTC 401
|||||
275 eutThrTyrPheGlyLeuPheGlnValHis 284
|||||
402 TAACCTACTTGGACTCTTTCAAGTTCAC 430
|||||

seq_name: gb_est1:AW917574

seq_documentation_block:
LOCUS AW917574 584 bp mRNA linear EST 25-MAY-2000
DEFINITION EST348878 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cDNA clone RGIERF49 5' end, mRNA sequence.
ACCESSION AW917574
VERSION AW917574.1 GI:8083328
KEYWORDS EST.

SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 584)
AUTHORS Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@ligr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M3 Reverse.

FEATURES
source
1. .584
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RGIERF49"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung,
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); site_1: EcoRI; site_2:
XhoI; Estimated insert size approx 1 kb"
BASE COUNT 108 a 188 c 147 g 138 t 3 others
ORIGIN

alignment_scores:
Quality: 652.00 Length: 135
Ratio: 5.015 Gaps: 0
Percent Similarity: 96.296 Percent Identity: 92.593

alignment_block:
US-09-245-198a-4 x AW917574 ..

Align seg 1/1 to: AW917574 from: 1 to: 584

150 ProArgProGlyGlnAspGlyAlaGlnAlaGlyValSerValSer 166
|||||
2 CCACAGCCAGCAGAGATGAGACAGCAGGCTGTGATGGAGACAGTGAG 51
|||||
166 rGlyTrpGluGluAlaArgIleAsnSerSerSerProLeuArgTyrAsn 183
|||||
52 TGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTGTGGCTATGAC 101
|||||
183 rGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyr 199
|||||
102 GCCAGATTGGGGAATTTACGGTCAATCAGGCTGGGCTTACTACTGTAC 151
|||||
200 CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLe 216
|||||
152 TGTCAGGTGCACTTTGATGAGGGAGGAGGAGCTGCTGGAAGCTGACTT 201
|||||
216 uLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaT 233
|||||
202 GCTGTGATGATGTGTGTGGCTTCGCTGCTGCTGGAATAATTCACGCA 251
|||||
233 hAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGly 249
|||||
252 CAGACGACAGCTCTCTCCGCGCCAGCTCCGCTGTGTGACAGTGTCTGG 301
|||||
250 LeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTr 266
|||||
302 CTGTGCTCTGTGGCGCAGGGTCTTCCCTCGAGATCCCTACCTCCCTCG 351
|||||

266 PALAHSLLeuLysAlaAlaProPheLeuThryrPheGlyLeuPheGlnY 283
|||||
352 GGCTCATCTTAAGGCGTCCCTTCTCTACCTACTTGGACTCTNNCAG 401
283 ALHIS 284
|||||
402 TTCAC 406
seq_name: gb_est2:BI906850
seq_documentation_block:
LOCUS BI906850 698 bp mRNA linear EST 16-OCT-2001
DEFINITION 603064633F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213480 5',
mRNA sequence.
ACCESSION BI906850
VERSION BI906850.1 GI:16169619
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 698)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1536 row: b column: 09
High quality sequence stop: 613.
Location/Qualifiers
1. 698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5213480"
/clone_lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 149 a 204 c 243 g 102 t
ORIGIN
alignment_scores:
Quality: 648.50 Length: 150
Ratio: 4.665 Gaps: 3
Percent Similarity: 92.667 Percent Identity: 92.000
alignment_block:
US-09-245-198a-4 x BI906850 ..
Align seg 1/1 to: BI906850 from: 1 to: 698
16 ProArSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnPr 32
11 |||||||
15 CCCCAGATCTCGGTCGCCGGATGGGGCGGTGAGGAGCAGCAGCC 64
32 OPFOAlaProMetAlaAlaArgArgSerGlnArgArgGlyArgArgG 49

|||||
65 CCCCCGCCCCATGCGCGCGCTCGAGCCAGAGCGGAGGGCGCGCG 114
49 lYgluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeu 65
115 GGAGCGGGGACACCGCTGTGTGTCCGCTGCGCTGGGCTGGGCTG 164
66 ALaLeuAlaCysLeuGlyLeuLeuAlaValAlaSerLeuGlySerAr 82
111 :||:|||||
165 GGTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
82 g.ALAserLeuSerAla...GlnGluProAlaGlnGluLeuValAla 97
215 GGGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
98 GlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerG 114
265 GAGGAGGACAGAGACCGCTGGAGACTGAATCCCGACAGAGAGAGCC 314
114 nAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaP 131
315 GGATCTGCGCGCTTCTCTGACCGACTAGTTGCGCTGCGACAGAGC 364
131 rOLysGlyArgLysThrArgAlaArgAlaAlaAlaAlaHisTyrGlu 147
365 CTAAAGCCCGGAAMACACGCGCTCGAAGAGCGATCCAGCCCATTA 414
148 VALHISProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
415 GTTCATCCAGCAGACTGTGACAGGAGGAGCGCAGGCA.....GATG 456
seq_name: gb_est2:BI908274
seq_documentation_block:
LOCUS BI908274 894 bp mRNA linear EST 16-OCT-2001
DEFINITION 603068526F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217367 5',
mRNA sequence.
ACCESSION BI908274
VERSION BI908274.1 GI:16171193
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 894)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1546 row: d column: 08
High quality sequence start: 5
High quality sequence stop: 460.
Location/Qualifiers
1. 894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5217367"
/clone_lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range


```

13 CCTCTCCCA...TCCCTCGGCTCCGGGATGGGGGGCCGCTAGGCAGGC 59
30 aGlnProFroAlaProMetAlaAlaArgArgSerGlnArgArgGlyA 47
60 ACAGCCCCCCCCCATGGCCCGCTGGAGCCAGAGCGGAGGGGGC 109
47 rGArgGlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeu 63
110 GCGGGGGGAGCGGGGACCGCCCTGCTGGCCCCCGCTGCTGAGCCTG 159
64 GlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuG 80
160 GGCCTGGGGCTGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 209
80 ySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuVal 97
210 GAGCTG.TTTATGTTGTCTGCCAGAGGCTTCTCAGAGGAGCTACAG 258
97 IaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluSer 113
259 CAGAGGACGCGCGGGAGGCCCTGAACTAATCCACAGACAGAGAAAC 308
114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerAl 130
309 CAGGATGTGTACTCTTCTTGGAAACAACATGACGCGCTCGAAGAGTGC 358
130 aProLysGluArgLysThrArgAlaArgArgAlaIleAlaAlaHisTyrC 147
359 TCCTAAAGGCGGAAAGCGCGGCTTCGCGGACCTAATTCAGGCCATTAG 408
147 LuValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
409 AGCTTCATCTCGGCCAGGACAGATGAGCACAAGCAGGTGTGATGGG 458
164 ThrValSerGlyTrp 168
459 ACAGTGAGTGGCTGG 473
seq_name: gb_est1:AA870722

```

SEQ. DOCUMENTATION BLOCK:	445 bp	mRNA	linear	EST 16-MAR-1998
LOCUS	AA870722			
DEFINITION	v225967.r1 Barstead stromal cell line MRLB8 Mus musculus cDNA clone IMAGE:1095324 5', mRNA sequence.			
ACCESSION	AA870722			
VERSION	AA870722.1	GI:2966167		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 445)			
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucada,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Teisling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.			
TITLE	The Washu-HHMI Mouse EST Project			
JOURNAL	Unpublished (1996)			
COMMENT	Contact: Marra M/Mouse EST Project			

FEATURES

Email: mouseest@atson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:601556
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 420.
Location/Qualifiers

```

source
1..445
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image:1095324"
/clone_id="Bartstead stromal cell line MPLR8"
/cell_line="C2C12 (undifferentiated)"
/lab_host="DH10B"
/notes="Vector: pT7.13-Pac (Pharmacia) with a modified
polylinker. Site.1: EcoRI; Site.2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGACATCGAGCGCGCGCCCTTTTCTTTTCTTTTCTTTT
3'], double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Source undifferentiated bovine culture cell line C2C12.
Library constructed by Bob Bartstead. The C2C12 cell line
(available from ATCC, catalog # CRL-1772) differentiates
rapidly, forming contractile myotubes and producing
characteristic muscle proteins."

```

```
alignment_scores:
  quality: 604.00      Length: 138
  Ratio: 4.576         Gaps: 0
  Percent Similarity: 95.652  Percent Identity: 86.957
```

```
alignment_block:
US-09-245-198A-4 x AA870722
```

Align seg 1/1 to: AA870722 from: 1 to: 445

[illegible]

Mon Jun 24 10:53:41 2002

OM of: US-09-245-198a-4 to: N_Geneseq_032802:* out_format : pfs

Date: Jun 22, 2002 2:32 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+r2n_model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US09245198/runat_21062002.184008.9995/app-query.fasta_1.632
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-EGAPOP=6.000 -EGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09245198 @CGN1_1_0 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-245-198a-4
Query length: 284
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 287.880000
```

score_list:

Sequence	Strd Orig	ZScore	Escore	len	! Documentation
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18600 +	1444.00	1871.03	7.2e-96	13	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV56000 +	1377.00	1783.63	5.3e-91	14	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49717 +	1366.00	1769.76	3.1e-90	13	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613 +	1265.00	1639.24	5.9e-83	12	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV04350 +	1265.00	1639.24	5.9e-83	12	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV03964 +	1066.00	1383.33	1.1e-68	8	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23424 +	1062.00	1376.94	2.4e-68	10	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18599 +	1020.00	1321.24	3.0e-65	11	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV23425 +	968.00	1258.09	1.0e-61	7	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV24140 +	127.00	152.58	3.7e	28	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49205 +	126.50	156.65	2.23	16	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV69289 +	124.50	149.40	5.6e	28	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAV25519 +	124.50	144.56	4.93.20	15	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV09686 +	124.00	145.84	8.93	3	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAV25519 +	124.00	113.91	536.08	15	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62160 +	123.50	135.74	32.60	11	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62176 +	123.50	115.68	427.27	11	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV46674 +	120.50	144.21	11.00	28	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV42224 +	117.50	137.51	25.99	35	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV79598 +	117.50	137.51	25.99	39	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV070738 +	117.00	137.51	26.00	39	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV99230 +	117.00	148.46	6.38	10	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV25214 +	117.00	148.46	6.38	10	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAV21473 +	116.00	147.61	7.12	9	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAV01620 +	115.00	144.00	11.31	12	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAV01318 +	115.00	143.54	11.99	13	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAV03062 +	115.00	143.53	12.00	13	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV4455 +	115.00	142.61	13.51	1	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV21793 +	115.00	133.60	42.94	4	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV41035 +	115.00	131.14	58.82	5	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV2729 +	115.00	127.97	88.34	6	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62133 +	114.50	123.33	160.14	12	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV05397 +	113.50	132.46	49.69	38	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAV78508 +	113.50	111.14	765.10	44	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAV80414 +	113.50	111.14	765.10	44	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAV29605 +	113.00	131.20	58.37	41	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAV29604 +	113.00	130.20	66.31	46	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV073500 +	113.00	124.94	130.26	84	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV63350 +	113.00	116.93	364.16	21	
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV63348 +	113.00	107.42	1.2e+03	63	

```
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV26029 + 112.50 140.19 18.42
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAV51445 - 112.50 136.66 29.00
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAV29515 + 112.50 132.26 51.00
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAV04352 + 112.50 131.91 53.32
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV4843 + 112.50 130.99 59.99
```

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18600

seq_documentation_block:

```
ID AAV18600 standard; cDNA; 1373 BP.
XX
XX AAV18600;
AC
XX 21-JUL-1998 (first entry)
DT
XX Homo sapiens tumour necrosis factor related ligand (TRELL) gene.
DE
XX TRELL; tumour necrosis factor related ligand; tnfr; treatment;
KM cancer; autoimmune disease; immune system; stimulation; suppression;
KM graft rejection; ds.
XX
XX Homo sapiens.
OS
XX
XX Key 1..852 Location/Qualifiers
FH CDS /tag= a
FT /note= "tumour necrosis factor related ligand"
FT
XX
XX W09805783-A1.
XX
XX 12-FEB-1998.
XX
XX 07-ANG-1997; 97WO-US13945.
XX
XX 18-MAR-1997; 97US-0040820.
XX 07-AUG-1996; 96US-0023541.
XX 18-OCT-1996; 96US-0028515.
XX
XX (BIOI ) BIOGEN INC.
PA (OYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
XX Browning JL, Chicheportiche Y;
XX
XX WPI: 1998-145619/13.
XX P-PSDB: AAM47525.
XX
XX Tumour necrosis factor related ligand - useful for, e.g. treating
XX cancer, auto-immune disease and immune responses to tissue grafts
XX
XX Claim 2: Pages 48-50; 69pp; English.
```

The sequence is that encoding human tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon- gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. The DNA sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective amounts of a vector, e.g. a virus comprising a gene encoding TRELL. It may also be of use in the preparation of prepare probes for screening natural/synthetic DNAs for TRELL-encoding sequences and for antisense therapy.

Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;

alignment_scores:

Quality: 1444.00 Length: 284
Ratio: 5.085 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-4 x AAV18600 ..

Align seg 1/1 to: AAV18600 from: 1 to: 1373

```
1 MetSerLeuLeuAspPheGluIleSerAlaArgAlaGluProLeuProAr 17
1 ATGTCATTGTTGACTTGAATTTCCGCCGCCGCTCCCTCCCTCCCGG 50
17 gSerLeuGlySerArgAspGlyGlyAlaValAlaArgAlaGluInProProA 34
51 ATCCCTCGGGTCCGGGATGGGGGGGGGTGAGCGAGCGACAGCCCGCG 100
34 laProMetAlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGlu 50
101 CCCCAGATGGCCGCCGTCGAGCAGAGCGGAGGGGGCGGGGGGAG 150
51 ProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuAlaLe 67
151 CCGGGCACCGCCTGCTGCTCCGCTCGCGCTGGCGCTGGCGCT 200
67 uAlaGlySerLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAlas 84
201 GCCCTGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
84 erLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluasp 100
251 CGCTGCGCGCCAGAGAGCTGCCAGAGAGAGCTGCTGCGAGAGAGGAC 300
101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAl 117
301 CAGGACCCGTCGGAAGCTGAATCCAGACAGAAAGCAGAGATCTGTC 350
117 aProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyAla 134
351 GCCTTCTCCAGACCACTAGTTCGGCTCGCGAGAGTGCACCTTAAGGCC 400
134 rGlySerThrArgAlaArgArgAlaAlaAlaAlaHisTyrGluValHisPro 150
401 GGAAGAACACGGGCTCGAAGAGCGATCGCAGCCATTATGAGTTCATCCA 450
151 ArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerG 167
451 CGACCTGGACAGAGAGCGAGCGAGAGGTGTGAGCGGACAGTGAAGTGG 500
167 yTrpGluGluAlaArgGlyLeaAsnSerSerSerProLeuArgTyrAsnArg 184
501 CTGGGAGGAGAGCCAGATCAACAGCTCCAGCCCTCTGCTGCTCAACCGCC 550
184 InIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrG 200
551 ACATCGGGAGATTATAGTCACCCGGGCTGGCTGCTACTGATCTGCTGT 600
201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 217
601 CAGGTGACACTTGTGATGAGGGGAGAGGCTGCTACCTGAACCTGGACTTGTCT 650
217 uValAspGlyValAlaLeuAlaLeuArgCysLeuGluGluPheSerAlaThrA 234
651 GGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
234 laAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeu 250
701 CGGCAGATTCTCGGGGCCCAAGCTCGGCTCTGCCAGGTGTCTGGGCTG 750
251 LeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrrAl 267
```

```
751 TTGGCCCTGGCGCCAGGGTCTCTCCCTGCGGATCCGACCTCCCTGGGCG 800
267 ahIseuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValH 284
801 CCATCTCAAGGCTGCCCTTCTCTCCTACTACTTGGAGCTTCCAGGTTTC 850
```

284 Is 284
851 AC 852

seq_name: /SIS1/gcdata/geneseq/geneseq-emb1/NA1999.DAT:AAx56000

seq_documentation_block:

ID AAX56000 standard; DNA: 1421 BP.

XX AAX56000;

DT 15-JUL-1999 (first entry)

DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.

KW Human: tumour necrosis factor: Apo-3 ligand; Lymphotoxin; apoptosis;

KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 92..841

FT /tag= a

FT /product= "Apo-3 ligand"

XX MO9919490-A1.

XX PD 22-APR-1999.

XX PF 09-OCT-1998; 98WO-US21407.

XX PR 17-DEC-1997; 97US-0069862.

XX PR 10-OCT-1997; 97US-0062037.

XX PA (GETH) GENENTECH INC.

XX PI Ashtkenazi AJ, Marsters SA, Pittl R;

XX DR WPI: 1999-287982/24.

XX DR P-PSDB: AAY09369.

XX PS New human Apo-3 ligand (a tumour necrosis factor) homologue

XX Claim 18; Fig 1; 74pp; English.

CC The present sequence encodes a human tumour necrosis factor (TNF) and

CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has

CC cytosolic activity. Apo-3 ligand can be used to induce apoptosis in

CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and

CC to induce JNK/SAPK-dependent responses in mammalian cells.

XX Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;

alignment_scores:

Quality: 1377.00 Length: 278
Ratio: 5.062 Gaps: 0
Percent Similarity: 97.842 Percent Identity: 97.842

alignment_block:

US-09-245-198a-4 x AAX56000 ..

Align seg 1/1 to: AAX56000 from: 1 to: 1421

```
7 GluIleSerAlaArgArgLeuProLeuProArgSerLeuGlySerArgas 23
|||||
```


5 GAGATCCCTGACCTGCAGCCACGCGTCCGCGATCCTCGGCTCCGGGA 54
XX
23 pGLyGlyAlaValArgGlnAlaGlnProProAlaProMetaAlaAlaArgA 40
XX
55 TGGGGGGGGGGTGGAGGAGGACAGCCGCCGCCCATGGCCGCCGCTC 104
XX
40 rGSerGlnArgArgArgGlyArgArgGlyGlnProGlyThrAlaLeuLeu 56
XX
105 GGAGCCAGAGGCGGAGGGGGGGGGGGAGCCGGGGACCGCCCTGCTG 154
XX
57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLeuGlyLeuLe 73
XX
155 GTCCCGCTGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGCT 204
XX
73 uLeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGln 90
XX
205 GCTGGCGCTGCTGCTGGGAGCCGGGCTGCTGCTGCTGCTGCTGCTG 254
XX
90 roAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 106
XX
255 CTGCCAGAGAGAGCTGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 304
XX
107 AspProGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 123
XX
305 AATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
XX
123 uValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgA 140
XX
355 AGTTCCGGCTCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 404
XX
140 rGAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 156
XX
405 GAGGAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 454
XX
157 AlAGlnAlaGlyAlaValAspGlyThrValSerGlyTrpGlnAlaArg 173
XX
455 GCGGAGGAG 504
XX
173 eAsnSerSerSerProLeuArgTrpArgAlaGlnGlnGlnGlnGlnGln 190
XX
505 CAACAGCTCCAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 554
XX
190 aThrArgAlaGlyLeuGlyTrpLeuGlyTrpCysGlnValHisPheAsp 206
XX
555 TCACCCGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
XX
207 GLyAlaAlaValTrpLeuGlyLeuAspLeuValAspGlyValLeuAl 223
XX
605 GGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 654
XX
223 aLeuArgCysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 240
XX
655 CTGGCGCTGCTGAGAGATTTCTGACCTGCGGCGAGATTTCTTCCG 704
XX
240 rGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgPro 256
XX
705 CCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
XX
257 SerSerLeuArgTrpLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnG 273
XX
755 TCCCTGCTGGGAGTCCGACCTCCCTGCGGCGCATCTCAAGGCTGCG 804
XX
273 oPheLeuThrTrpPheGlyLeuPheGlnValHis 284
XX
805 CTTCCTCACCCTACTTCGAGCTTCCAGGTTTCAC 838
XX

seq_name: /STD1/9c9data/geneseq/geneseq_emb1/NA2000.DAT:AAA49717

seq_documentation_block:

ID AAA49717 standard: cDNA: 1353 BP.

AC AAA49717;
XX

DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 cDNA clone DNA30879-1152.
XX
KW PRO207; human; antitumour; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..807 /*tag= a
FT sig_peptide 58..177 /*tag= b
FT mat_peptide 178..804 /*tag= c
FT
XX
PN WO200037638-A2.
XX
PD 29-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28565.
XX
PF 22-DEC-1998; 98US-0113296.
XX
PR 08-MAR-1999; 99WO-US05028.
XX
PR 21-APR-1999; 99US-0130232.
XX
PR 28-APR-1999; 99US-0131445.
XX
PR 14-MAY-1999; 99US-0134287.
XX
PR 20-JUL-1999; 99US-0144758.
XX
PR 26-JUL-1999; 99US-0145698.
XX
PR 15-SEP-1999; 99WO-US21090.
XX
PR 15-SEP-1999; 99WO-US21547.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA:
PI Napier MA, Pitti RM, Wood WI;
XX
DR WPI: 2000-442668/38.
XX
DR P-PSDB: AAY95338.
XX
PT Novel composition to inhibit neoplastic cell growth or for treating
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or
PT PRO866
XX
PS Claim 20; Fig 3; 172pp; English.
XX
CC The present sequence is that of cDNA clone DNA30879-1152
CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumour necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a
CC foetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukaemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

alignment_scores:

Quality: 1366.00 Length: 268
Ratio: 5.097 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-245-198a-4 x AAA49717 ..
Align seg 1/1 to: AAA49717 from: 1 to: 1353

```

17 ArgSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProPr 33
1 CGATCCCTCGCGGATGGGGGCGGTGAGGACAGGACAGACCCCC
33 GAlaProMetAlaAlaArgArgSerGlnArgArgArgGlyArgGlyG 50
51 CCCCCCATGGCCCGCTGGAGCCAGAGCGGAGGGGCGCCCGGGGG 100
50 LProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAla 66
101 AGCCGGGACCCCGCTGCTGCTCCGCTGCGCTGGGCTGGGCTGGGG 150
67 LeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAl 83
151 CTGGCCTGCGCTGCGCTGCTGCGCTGCGCTGCTGCTGCGCTGCGG 200
83 eSerLeuSerAlaGlnGlnProAlaGlnGlnLeuValAlaGlnGln 100
201 ATCGCTGCTCGCCAGAGAGCTGCGCCAGAGAGAGCTGCTGCGAGAG 250
100 sPGlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspPro 116
251 ACCAGAGACCCCTGCGACTGATCCCGACAGAGAGAGAGAGAGATCCT 300
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysG 133
301 GCGCCTTCTCCAGACGACTACTGCGCTGCGCTGCGAGAGTGAAG 350
133 YArgIysThrArgAlaArgArgAlaIleAlaIleHisThrGlyIleValHis 150
351 CCGGAAACACGGGCTCGAAGAGCGATCGACCCCATTTATGAAATTCAATC 400
150 rGARProGlyGlnAspGlyAlaGlnAlaGlnValAspGlyThrValSer 166
401 CACGACCTGGACAGGAGCGAGCGAGCGAGGTGTGGACGGAGACGACT 450
167 GlyTrpGlnGlnAlaArgIleAsnSerSerProLeuArgTyrAsnAr 183
451 GCGTGGAGAGAGCCAGATCAACAGCTCCAGCCCTGCTGCTACAAACG 500
183 gGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrC 200
501 CCAGATCGGGGAGTTATATGTCACCCGGGCTGGGCTGCTACTACT 550
200 ySGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeu 216
551 GTCAGGTGCACTTGTATGAGGGAGAGCGTCTACTGTAAGCTGGACT 600
217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGlnPheSerAlaTh 233
601 CTGGTGGAGAGGTGTGCTGCGCTGCGCTGCGAGAAATTTCAACCCAC 650
233 rAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
651 TCGGGGAGATCTCCCTGGGCGCCAGCTCCGCTGCTCCAGGTGTGGGC 700
250 eUleuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp 266
701 TGTGTGGCCCTGGGCGCAGGCTCTCTCCGATCCGACACCTCCCTGG 750
267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVa 283
751 GCCCATCTCAAGGCTGCCCTTCTCACTACTTGGAGACTCTTCAGAGT 800
283 His 284
1111
801 TCAC 804

```

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613
seq_documentation_block:
ID AAV47613 standard; CDNA: 1236 BP.

```

XX AC AAV47613;
XX DT 27-OCT-1998 (first entry)
XX DE TNF related endothelium proliferative agent gene.
XX KW ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
XX KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..750
XX FT /*tag= a
XX FT /*product= "TREPA"
XX PN M09835061-A2.
XX PD 13-AUG-1998.
XX PF 12-FEB-1998; 98MO-US02859.
XX PR 10-FEB-1998; 98US-0021706.
XX PR 12-FEB-1997; 97US-0798692.
XX PA (ABBO ) ABBOTT LAB.
XX PI WILLEY SR.
XX DR WPI: 1998-447255/38.
XX DR P-PSDB: AAM29745.
XX PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
XX PT treatment of autoimmune disease, tumours and inflammation
XX PS Claim 11: Page 123-4; 142pp; English.
XX CC The TNF-related endothelium proliferative agent (TREPA), or its
XX CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
XX CC promote wound healing or tissue grafting, by promoting vascularisation,
XX CC also to induce apoptosis for treating cancer and eliminating autoreactive
XX CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
XX CC TREPA peptides can also be used to target cytotoxic agents or for
XX CC affinity isolation of the corresponding receptor, the nucleic acid for
XX CC which can be used to transform tumour cells to render them more
XX CC responsive to TREPA and to screen for TREPA mimics.
XX CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
XX CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
XX CC vascularisation), inflammation or a wide range of autoimmune conditions,
XX CC conditions involving abnormal stimulation of epithelial cells (e.g.
XX CC atherosclerosis), for birth control (inhibiting ovulation and placental
XX CC formation) or other angiogenic conditions (e.g. ulcers).
XX SO Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

```

alignment_scores:
Quality: 1265.00 Length: 249
Ratio: 5.080 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.598

alignment_block:
US-09-245-198a-4 x AAV47613 ..

Align seg 1/1 to: AAV47613 from: 1 to: 1236

36 MetAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGlnProG 52

```

|||||
1 ATGGCCCCCGTCGAGCCAGAAAGCGAGGGGGCCGGGGGAGCCGGG
52 yTtRAlAlaLeuValProLeuAlaLeuGlyLeuAlaLeuAlaLeuAlaC 69
51 CACCGCCCTCTGCTGCTCCGCTCGGCTGGGCTGGGCTGGGCTGGCT 100
69 ySLeugLyLeuLeuAlaValaValSerLeuGlySerArGAlaSerLeu 85
101 GCCGCGGCTCTGCTGCTGGCTGGTCAgTTTGGGAGCCGGGCAATGCTG 150
86 SerAlaGInGluProAlaGInGluGluLeuValaLaGluGluAspGlnAs 102
151 TCGGCCAGAGAGCTGCCAGAGAGAGCTGTGCGAGAGAGAGAGAGAGAG 200
102 pProSerGluLeuAsnProGlnThrGluGluGluSerGlnAspProAla 119
201 CCCGTCGGAAGTGAATCCCAAGACAGAAAGAGAGAGATCTGGCTCTT 250
119 heLeuAsnArGLeuValaArgProArGArGSerAlaProLyGlyArgLy 135
251 TCCGGAACGACACTAGTCCGGCTCGAAGAAGTGCACCTAAAGCCGGA 300
136 ThrArGAlaArGArGAlaAlaAlaAlaHisTyrGluValHisProAr 152
301 ACACGGGCTCGAAGAGCGATCGACCATATGAATGATCATCCACGAC 350
152 oGlyGInAspGlyAlaGInAlaGlyValaAspGlyThrValSerGlyTr 169
351 TGGACAGAGAGAGAGCGAGCGAGGTGTGACGGGAGAGTAGTGCGGG 400
169 luGluAlaArGlyLeuAsnSerSerSerProLeuArGlyrAsnArGly 185
401 AGGAAGCGCAATCAACAGAGCTCCAGCCCTCGCCCTACAAACCGCA 450
186 GlyGluPheIleValThrArGAlaGlyLeuTyrTyrLeuTyrGlyNa 202
451 GGGGAGTTTATGATCACCGGGCTGCTACTACCTGTACTGTACGT 500
202 lHisPheAspGluGlyLyAlaValTyrLeuLyLeuAspLeuVala 219
501 GCACCTTGAAGAGGGAGAGCTGTCTACAGTGAAGCTGAGCTGTGTTG 550
219 spGlyValLeuAlaLeuArGlySerLeuGluGluPheSerAlaThrAla 235
551 ATGGTGTGCTGGCCCTGCTGCTGAGAGATTCACGACACGCGCG 600
236 SerSerLeuGlyProGlnLeuArGLeuGlyGlnValSerGlyLeuAl 252
601 AGTTCCCTCGGGCCCGAGCTCCGCTGTGCCAGGTGTGGGTGTGGC 650
252 aleuArGProGlySerSerLeuArGlyLeuArGThrLeuProThrAla 269
651 CCGGCGCCAGGGCTCTCTCGGAGATCCGACCCCTCCCTGGGCCATC 700
269 eutysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
701 TCAAGGCTGCCCTTCTCTCACCCTTCCGAGCTTCCAGGTTCCAC 747

seq_name: /SIDSI/gcgsdata/geneseq/geneseq-emb1/NA2001A.DAT.AAD04350
seq_documentation_block:
ID AAD04350 standard; CDNA; 1236 BP.
XX
AC AAD04350;
XX
DT 04-JUN-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnery; ss.

```

```

XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 1..750
FT CDS /tag=a
FT /product="Human TREPA (TNF related endothelium
FT proliferative agent)"
XX
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-0105343.
XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
XX (ABD0 ) ABBOTT LAB.
XX
XX WILEY SR;
XX
XX WPI: 2001-280760/29.
XX
XX P-PSDB: AAE00891.
XX
XX Inducing angiogenesis in mammal at desired sites for promoting wound
XX healing, by administering soluble fragment of extracellular domain of
XX tumor necrosis factor related endothelium proliferative agent protein
XX
XX Example 2; Column 73-74; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,
XX particularly members of tumour necrosis factor (TNF) family molecules
XX designated as TREPA (TNF related endothelium proliferative agent).
XX Soluble biologically active TREPA are used to treat TREPA-associated
XX diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX in human for promoting wound healing and for vascularising grafted
XX tissue for successful grafting and to promote tissue grafts.
XX The present sequence is a cDNA clone ID #690050 encoding human TREPA.
XX
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
XX
XX
XX alignment_scores:
XX Quality: 1265.00 Length: 249
XX Ratio: 5.080 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 99.598
XX
XX alignment_block:
XX US-09-245-198a-4 x AAD04350 ..
XX
XX Align seq 1/1 to: AAD04350 from: 1 to: 1236
XX
XX 36 MetaAlaAlaArGArGSerGlnArGArGlyrArGArgGlyGluProG 52
XX 1 ATGGCCCCCGTCGAGCCAGAAAGCGAGGGGGCCGGGGGAGCGGG 50
XX 52 yTtRAlAlaLeuValProLeuAlaLeuGlyLeuAlaLeuAlaLeuAlaC 69
XX 1 ATGGCCCCCGTCGAGCCAGAAAGCGAGGGGGCCGGGGGAGCGGG 50
XX 51 CACCGCCCTCTGCTGCTCCGCTCGGCTGGGCTGGGCTGGGCTGGCT 100
XX 69 ySLeugLyLeuLeuAlaValaValSerLeuGlySerArGAlaSerLeu 85
XX 101 GCCGCGGCTCTGCTGCTGGCTGGTCAgTTTGGGAGCCGGGCAATGCTG 150
XX 86 SerAlaGInGluProAlaGInGluGluLeuValaLaGluGluAspGlnAs 102
XX 151 TCGGCCAGAGAGCTGCCAGAGAGAGCTGTGCGAGAGAGAGAGAGAG 200
XX 102 pProSerGluLeuAsnProGlnThrGluGluGluSerGlnAspProAla 119
XX 201 CCCGTCGGAAGTGAATCCCAAGACAGAAAGAGAGAGATCTGGCTCTT 250

```

```

119  heLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
    |||||||
251  TCCTGAACCGACTAGTTGGCTCGAAGACGACCTAAAGCGCGGAAA 300
    |||||||
136  ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgPr 152
    |||||||
301  ACAGCGGCTCGAAGAGCGATCGACGCCATTATGAAATTCAATCCAGACC 350
    |||||||
152  oGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPG 169
    |||||||
351  TGGACGAGGAGGAGCGACGAGGTGTGACGGGACAGTGAAGTGGCTGGG 400
    |||||||
169  IuGluValArgIleAsnSerSerProLeuArgTyrAsnArgGlnIle 185
    |||||||
401  AGGAAGCCAGATCAACAGCTCCAGCCCTGCGGTACACCGCCAGATC 450
    |||||||
186  GlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVa 202
    |||||||
451  GGGGAGTTTATAGTCACCCGGGCTGGCTTACTACTGTCTACTGTCAAGT 500
    |||||||
202  IHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValA 219
    |||||||
501  GCACCTTATGAGGGAGAGGCTGTCTACCTGAAGCTGAGCTTGGCTGG 550
    |||||||
219  spGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
    |||||||
551  ATGGGTCTCTGGCCCTGGCTGGCTCGAGAGAAATTCACGCCCTCGGCG 600
    |||||||
236  SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAl 252
    |||||||
601  AGTTCCCTCGGGCCCCACCTCGGCTCGCAGGAGTGTGGGCTGTGGC 650
    |||||||
252  aLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisL 269
    |||||||
651  CCTGGGCGGACAGGGTCTCTCCCTGGGATCCGACCTCCCTCGGCGCCATC 700
    |||||||
269  euLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
    |||||||
701  TCAAGGCTGCCCCCTTCTCAACCTACTCTCGGACTCTTCACAGGTTTCAC 747

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA503964
seq_documentation_block:
ID   AA503964 standard; DNA; 898 BP.
XX
AC   AA503964;

26-SEP-2001 (first entry)

Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
XX
XX   TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis;
    ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
    retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
    rubecosis; uveitis; macular degeneration; arthritis; Rheumatism; ds;
    corneal graft neovascularisation; psoriasis; metastatic condition;
    malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
    preneoplastic condition; myocardial angiogenesis; wound granulation;
    scleroderma; vascular adhesion; telangiectasis; ischemia; human;
    atherosclerotic plaque neovascularisation; coronary atherosclerosis;
    peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
    fusion protein.
XX
XX   Homo sapiens.
OS   Synthetic.
XX
XX   Key Location/Qualifiers
FH   52..873
FT   CDS /tag= a
FT   /product= "fusion protein comprising a growth hormone
FT   leader, a leucine zipper multimerisation
FT   domain, and human TWEAK extracellular

```

```

FT   domain"
XX
XX   MO200145730-A2.
PN
XX
XX   28-JUN-2001.
PD
XX
XX   19-DEC-2000; 2000MO-US34755.
PF
XX
XX   20-MAY-1999; 99US-0172878.
PR   10-DEC-2000; 2000US-0203347.
XX
XX   (IMMV ) IMMUNEX CORP.
XX
XX   Willey SR.
PI
XX
XX   WPI: 2001-417975/44.
DR
XX
XX   P-PSDB; AAU03459.
XX
XX   Modulating angiogenesis in a mammal for treating diseases mediated by
XX   angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX   peripheral tissue, by administering antagonist or agonist of TWEAK
XX   receptor
XX
XX   Example 1: Page 39-40; 46pp: English.
XX
XX   The sequence represents a DNA from the expression vector
XX   pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth
XX   hormone leader, a leucine zipper multimerisation domain, and the
XX   extracellular domain of human TWEAK. The fusion protein was used in
XX   the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX   from a COS cell human cDNA library. The TWEAK protein is a
XX   member of the tumour necrosis factor (TNF) family and induces
XX   angiogenesis. TWEAKR may therefore be used to screen for and develop
XX   TWEAK agonists and antagonists for the modulation of angiogenesis, to be
XX   used in the treatment and diagnosis of human disease. The disorders
XX   mediated by angiogenesis include ocular disorders characterised by ocular
XX   neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX   retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX   rubecosis, uveitis, macular degeneration and corneal graft
XX   neovascularisation, and inflammatory diseases such as arthritis,
XX   rheumatism and psoriasis. Other treatable diseases include malignant and
XX   metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX   preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX   scleroderma, vascular adhesions, atherosclerotic plaque
XX   neovascularisation, telangiectasia, wound granulation, coronary
XX   atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
XX   Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:
SQ

alignment_scores:
    Quality: 1066.00      Length: 207
    Ratio: 5.150          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-245-198A-4 x AA503964 ..
Align seg 1/1 to: AA503964 from: 1 to: 898

78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluCl 94
|||||
250 AGTTGGGAGAGCCGGGAGATCGCTGCGCCCGAGGAGCTCGCCAGAGCA 299
|||||
94 IuGluValAlaGluGluAspGluAspProSerGlyLeuAsnProGlnTrng 111
|||||
300 GCTGGTGGCAGAGGAGGACGACGACCGCTCGGAATGAAATCCCGAGACG 349
|||||
111 IuGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArg 127
|||||
350 AAGAAGCCAGATCTCGGCTTCTCGAAGACGACTGTTGGCTCGC 399
|||||
128 ArgSerAlaProLysGlyArgLysThrArgAlaArgArgAlaIleAlaAl 144

```

```

400 AGAAGTGCACCTAAAGCCGGAACACGGGCTGCAAGAGCGATCGCAGC 449
144 AHIISTYGLIUAHISPROARGPROGLYGLINASPGLYVALAGINLAGY 161
450 CCATTATGAACTTCATCCAGCCTCGGACAGACGAGCGAGCGAGGCTG 499
161 AIAISPGLYTHRVAlSerGIYTRPGIUGIUAIAARGLIeASerSer 177
500 TGGACGGGACAGTGAAGTGGCTGGGAGGAGGACCAATCAACAGCTCCAGC 549
178 PROLeuArGIYrAsnArGIInIleGIYGLIuPheIleValThrArGAlaGI 194
550 CCTTGCGGCTGACAACCGCAGATCGGGAGGAGTTATAGCACCCGGGCTGG 599
194 YLeuTYrTYrLeuTYrCYSGIInValHISPhASpGLIUGIYsAlaValT 211
600 GCTGTACTACTGCTGACTGTCAAGTGCACCTTTGATGAGGGAGGCTGTCT 649
211 YrLeuYsLeuAspLeuLeuValAspGIYValLeuAlaLeuArGIYsLeu 227
650 ACCTGAAGCTGACTGTGCTGGATGGTGTGCTGGCCCTGCGCTGCTG 699
228 GLIUGIuPheSerAlaThrAlaAlaSerSerLeuGIYProGIInLeuArGI 244
700 GAGGAATTCAGCAGCAGTGGCGGAGTTCCCTCGGGCCCGCAGTCCGCT 749
244 uCYSGIInValSerGIYLeuLeuAlaLeuArGIYProGIYSerSerLeuArGI 261
750 CTGTCAGAGTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 799
261 IeArGIYrLeuProTrPAlaHISLeuYsAlaAlaProPheLeuThrTYr 277
800 TCGCGACCTCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTAC 849
278 PheGIYLeuPheGIInValHIS 284
850 TTCGAGCTCTTCAGGTTTCAC 870

```

seq_name: /std1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAx23424

seq_documentation_block:

ID AAX23424 standard; DNA; 1030 BP.

XX AAX23424;

DT 18-JUN-1999 (first entry)

XX Human TNRL3 DNA.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KM developmental abnormality; gestational abnormality; prostate cancer;
 KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KM apoptosis; human; ss.

XX Homo sapiens.

XX key Location/Qualifiers

FT CDS 1..627 /tag= a

XX MO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98MO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNIT) UNIV WASHINGTON.

XX

```

PI Chaudhary PM:
XX
DR WPI: 1999-205191/17.
DR P-PSDB; AAM93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX
PS Example VII: Fig 13A; 156bp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:

```

alignment_scores: Quality: 1062.00 Length: 208
 Ratio: 5.130 Gaps: 0
 Percent Similarity: 99.519 Percent Identity: 99.519

alignment_block: US-09-245-198A-4 x AAX23424 ..

Align seg 1/1 to: AAX23424 from: 1 to: 1030

```

77 ValSerLeuGIYSerArGIAlaSerLeuSerAlaGIInGIuProAlaGIInGI 93
1 GTCAAGTTTGGGGAGCGCGGATCGCTGTCCGCCAGAGAGCTGCCAGCA 50
93 uGIInValAlaGIInGIuAspGIInAspProSerGIInLeuAsnProGIInT 110
51 GGAGCTGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 100
110 hrgIUGIuSerGIInAspProAlaProPheLeuAsnArGIYValArGIPro 126
101 CAGAAAGAAAGCCAGATCTCGGCTTCTCGAACCAGACTGTGCGGCT 150
127 ArGIYrSerAlaProLYSGIYArGIYrThrArGIAlaArGIAlaIleAl 143
151 CGCAGAAAGTGCACCTAAAGCGCGGAAACACGGGCTCGAAGAGCGATCGC 200
143 AAlaHISrYrGIInValHISProArGIProGLIUGIInAspGIYValAGInAG 160
201 AGCCATTATGAACTTCATCCAGCCTGGAGAGGAGGAGGAGGAGGAGGAG 250
160 YrValAspGIYThrValSerGIYTRPGIUGIUAIAARGLIeASerSer 176
251 GTGTGAGCGGAGAGTGAAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAG 300
177 SerProLeuArGIYrAsnArGIInIleGIYGLIuPheIleValThrArGIAl 193
301 AGCCCTGTGCGCTACTACCGCAGATCGGGAGGAGTTATAGTACACCGGCG 350

```

```

193 acgyleuTYrTYrleuTYrCysGlnValHisPheaspGluGlyLysAlaVal 210
|||||
351 TGGGCTCTACTACTGTACTGTACAGTGCACCTTGATGAGGGAAAGCGTG 400
|||||
210 aTYrleuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCys 226
|||||
401 TCTACCTGAAGCTGGACCTGTGGTGAGTGCTGTGGCCCTGCGCTTC 450
|||||
227 LeuGluGlnPheSerAlaThrAlaHisSerLeuGlyProGlnLeuArg 243
|||||
451 CTGGAGGAATTTCTCAGCCACTGGGCACTTCCCTGGGCCCCAGCTCCG 500
|||||
243 gLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeu 260
|||||
501 CCTGTGCCAGTGCTGTGGCTGTGGCCCTGGGCGGAGGCTCCCTCCG 550
|||||
260 rGllLeuArgThrLeuProThrAlaHisLeuValAlaAlaProheLeuThr 276
|||||
551 GGATCGGACCTCCCTGGGCGGCACTTCAAGGCTGCCCTTCCTCACCC 600
|||||
277 TYrPheGlyLeuPheGlnValHis 284
|||||
601 TACTTCGGACTCTTCCAGGTTTCAAC 624
|||||

```

seq_name: /SIDSL/gcdata/geneseq/geneseq-emb1/NA1998.DAT:AAV18599

seq_documentation_block:

ID AAV18599 standard; cDNA: 1168 BP.

XX AAV18599;

DT 21-JUL-1998 (first entry)

XX Mus musculus tumour necrosis factor related ligand (TRELL) gene.

XX TRELL: tumour necrosis factor related ligand; tnfr; treatment;

XX cancer; autoimmune disease; immune system; stimulation; suppression;

XX graft rejection; ds.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 2..679 /*tag= a /note= "tumour necrosis factor related ligand"

XX MO9805783-A1.

XX 12-FEB-1998.

XX 07-AUG-1997; 97WO-US13945.

XX 18-MAR-1997; 97US-0040820.

XX 07-AUG-1996; 96US-0023541.

XX 18-OCT-1996; 96US-0028515.

XX (BIOJ) BIOGEN INC.

XX (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicheportliche Y;

XX WPI: 1998-145619/13.

XX P-PSDB: AAM47524.

XX Tumour necrosis factor related ligand - useful for, e.g. treating

XX cancer, auto-immune disease and immune responses to tissue grafts

XX Claim 2; Pages 45-46; 69pp: English.

XX The sequence is that encoding mouse tumour necrosis factor related

XX ligand (TRELL). TRELL or active fragments can be included with a

XX carrier in pharmaceutical compositions to treat cancer, autoimmune

CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. The DNA sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.

SO Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;

alignment_scores:

Quality: 1020.00 Length: 224
 Ratio: 4.744 Gaps: 0
 Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198A-4 x AAV18599

Align seg 1/1 to: AAV18599 from: 1 to: 1168

```

61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValVa 77
|||||
5 CTGAGCGCTGGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 54
|||||
77 IserLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94
|||||
55 CAGCGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAG 104
|||||
94 IuLeuValAlaGluGlnAlaAspGlnAspProSerGlnuLeuAsnProGlnThr 110
|||||
105 AGCTGACACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 154
|||||
111 GluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAr 127
|||||
155 GAGAAAGCAGAGATGTGTACTTCTTGAACAACACTAGTCGGGCTCG 204
|||||
127 gArgSerAlaProLysGlyArgLysThrArgAlaArgArgAlaIleAla 144
|||||
205 AAGAAAGTCTCTAAAGCGGAGGCGGCGGCTCGGCGGCTATTTGCG 254
|||||
144 IahIstYrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
|||||
255 CCCATTATAGAGTTCATCTCGGCCAGACAGATGAGACACAGAGAGT 304
|||||
161 ValAspGlyThrValSerGlyTyrGlnGluAlaArgIleAsnSerSerse 177
|||||
305 GTGGATGGGAGACAGTGTGTGGGAGAGACCAAAATCAACAGCTCCAG 354
|||||
177 rProLeuArgTyrAsnArgGlnIleGlyLuhPheIleValThrArgAlaG 194
|||||
355 CCTGTGGCTAGACCGCCAGATTGGGAATTTACAGTCATCAGAGCTG 404
|||||
194 IYleuTYrTYrleuTYrCysGlnValHisPheaspGluGlyLysAlaVal 210
|||||
405 GGCTCTACTACTGTACTGTACAGTGCACCTTGATGAGGAAAGCGTGC 454
|||||
211 TYrleuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCysLe 227
|||||
455 TACCTGAAGCTGACTGTGTGGAGGCTGTGGCTGGCTGGCTGGCTGG 504
|||||
227 uGluGluPheSerAlaThrAlaHisSerLeuGlyProGlnLeuArgL 244
|||||
505 GGAAGAATTTCTCAGCCACAGCAAGACTCTCTCTGGGCCCGAGCTCGTT 554
|||||

```

```
244 eucysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260
|||||
555 TGTCGACAGGTGTCTGGGCTGTGGCGGCGGACAGGCTCTCCCTTCGG 604
261 ILeArgThrLeuProTrrPAlaHisLeuLysAlaAlaProPheLeuThr 277
|||||
605 ATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTCTTCTTAACCTA 654
277 rPhGlyLeuPheGlnValHis 284
|||||
655 CTTTGACCTCTTCAAGTTCAC 676
seq_name: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:AAx23425
```

```
seq_documentation_block:
ID   AAX23425 Standard; DNA; 701 BP.
```

```
AC   AAX23425;
```

```
DE   18-JUN-1999 (first entry)
```

```
XX   Mouse TNRL3 DNA.
```

```
XX   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; mouse; ss.
```

```
XX   Mus sp.
```

```
XX   Key      Location/Qualifiers
FT   CDS      1..636
FT           /tag= a
FT           /product= "TNRL3"
```

```
XX   WO9911791-A2.
```

```
XX   11-MAR-1999.
```

```
XX   04-SEP-1998; 98WO-US18393.
```

```
XX   05-SEP-1997; 97US-0924634.
```

```
XX   (UNITW ) UNIV WASHINGTON.
```

```
XX   Chaudhary PM;
```

```
XX   WPI: 1999-205191/17.
```

```
XX   P-PSDB: AAM93591.
```

```
XX   New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX   useful for diagnosis and treatment of prostate cancer and
XX   developmental or gestational abnormalities
```

```
XX   Example VII: Flg 13B; 156bp; English.
```

```
XX   This invention describes isolated Tumor Necrosis Factor (TNF) family
XX   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX   their active fragments. APO4 is useful for diagnosing prostate cancer
XX   by determining levels of APO4 in an individual. Prostate cancer can also
XX   be treated using APO4 selective binding agents linked to a therapeutic
XX   moiety. APO4 polypeptides are also useful for identifying selective
XX   binding agents, useful in diagnosis/treatment of disease by binding of
XX   agents to the polypeptide/active fragment which is extracellular, or
XX   expressed on the cell surface. The binding is preferably performed in
XX   vivo. APO4 polypeptides/active fragments are also useful for screening
XX   for agonists and antagonists by binding and observing the changer in APO4
XX   activity. Effective pharmacological agents useful in diagnosis or
XX   treatment of disease are also identified using APO4 polypeptides/active
XX   fragments and APO4 signal transducer molecules that specifically interact
```

```
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
```

```
XX   Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;
```

```
alignment_scores:
```

```
Quality: 968.00
```

```
Ratio: 4.768
```

```
Gaps: 211
```

```
Percent Similarity: 96.209 Percent Identity: 89.100
```

```
alignment_block:
```

```
US-09-245-198a-4 x AAX23425 ..
```

```
Align seg 1/1 to: AAX23425 from: 1 to: 701
```

```
74 LeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGln 90
|||||
1 CTGGTCGTGTGTCAGCTGGGAGCTGGGCAACGCTGCTGCGCCAGAGCC 50
90 cAlaGlnGlnGlnLeuValAlaGlnGlnAlaGlnAspGlnAspSerGlnLeuA 107
|||||
51 TTTCAGAGAGAGAGCTGACACAGAGAGACCGCGGAGACCCCTCACTGA 100
|||||
107 snProGlnThrGlnGlnSerGlnAspProAlaProPheLeuAsnArgLeu 123
|||||
101 ATCCCGACAGAGAGAGAAAGCCAGAGATGTGTACTTCTTGAACAACATA 150
124 ValArgProArgSerArgAlaProLysGlyArgLysThrArgAlaArgAr 140
|||||
151 GTCCGCGCTCGAAGAGAGCTCTTAAAGCGGAGGAGCGGCTCCCGG 200
|||||
140 gAlaAlaAlaAlaHisTrrGlnValHisProArgProGlyGlnAspGlyA 157
|||||
201 AGCTATTGCACGCCCATTTATGAGTTTATCTCCGCGCCAGAGAGATGAG 250
|||||
157 laGlnAlaGlyValAspGlyThrValSerGlyTrrGlnGlnAlaArgIle 173
|||||
251 CACAGCAGAGGTGTGATGGAGACAGTGAAGTGGCTGGAAAGACCAAAATC 300
|||||
174 AsnSerSerSerProLeuArgTrrAsnArgGlnIleGlyGlnPheIleA 190
|||||
301 AACAGCTCACGCCCTCTGCGCTACGACGCCAGATTGGGAATTACAGT 350
190 lThrArgAlaGlyLeuTrrTrrLeuTrrCysGlnValHisPheAspGln 207
|||||
351 CATCAGGCGTGGGCTCTACTACCTGCTACTGTCAGGTGCACCTTTATATAG 400
207 lLysAlaValAlaTrrLeuLysLeuAspLeuValAspGlyValLeuAla 223
|||||
401 GAAAGCGTGTCTACTGAGCTGACTGCTGGAGACGGTGTCTGGCC 450
224 LeuArgGlyLeuGlnGlnPheSerAlaThrAlaAlaSerLeuGlyPrr 240
|||||
451 CTGGCTGCTCGAAGAAATTCCTCAGCCACAGCACAGGCTCTCTGGGCC 500
240 oGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGlys 257
|||||
501 CCAGCTCCGTTTGTGCGCAGGTGTCTGGGCTTTGGCCCTGGCGGACAGGT 550
257 eTrSerLeuArgIleArgThrLeuProTrrPAlaHisLeuLysAlaAlaPro 273
|||||
551 CTTCCTTCCTCCGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCC 600
274 PheLeuThrTrrPheGlyLeuPheGlnValHis 284
|||||
601 TTCTTAACCTTACCTTGGAGCTCTTCAAGTTTCAC 633
seq_name: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV24140
```


277 TyphedglyleupheglnVal 283
 |||||
 2645 TAAACGTGGCTTTTGACATT 2665

seq_name: /STD1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA49205

seq_documentation_block:
 ID AAA49205 standard; cDNA; 1660 BP.

AC AAA49205;

DT 12-DEC-2000 (first entry)

DE Corn putative lecithin:cholesterol acyltransferase gene #5.

KM Corn: lecithin:cholesterol acyltransferase; phytoosterol;
 XX phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.

OS Zea mays.

Key Location/Qualifiers

FT CDS 77..1396

FT /tag= a

FT /product= "lecithin:cholesterol acyltransferase"

PN MO200032791-A2.

PD 08-JUN-2000.

PF 02-DEC-1999; 99WO-US28586.

PR 03-DEC-1998; 98US-0110782.

PA (DUP0) DU PONT DE NEMOURS & CO E. I.

PI Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;

DR WPI: 2000-412337/35.

XX P-PSDB; AAB01210.

PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase

PT enzyme useful for producing transgenic plants and for producing

PT antibodies specific to which is useful for screening cDNA expression

PT libraries -

XX Claim 2; Page 43-44; 49pp; English.

XX The present sequence is a putative coding sequence for a corn

XX lecithin:cholesterol acyltransferase (also known as

XX phosphatidylcholine-sterol O-acyltransferase). This enzyme is found

XX associated with high-density lipoproteins and Apolipoprotein AI and -D.

XX The gene and protein can be used to produce transgenic plants which have

XX increased lipid metabolism and membrane fluidity, and therefore increased

XX resistance to heat and/or cold shock, to alter the content of phytoosterol

XX or lecithin in grains and to identify potential herbicides.

XX Sequence 1660 BP; 361 A; 475 C; 499 G; 325 T; 0 other;

alignment_scores:

Quality: 126.50 Length: 319

Ratio: 0.866 Gaps: 17

Percent Similarity: 45.768 Percent Identity: 25.392

alignment_block:

US-09-245-198A-4 x AAA49205 ..

Align seg 1/1 to: AAA49205 from: 1 to: 1660

11 ArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaVal 27

394 CAAAGTACCGAAGCCTCGCCGG.....CGTCGAGAGCGGAGTGCACA 437

27 1.....ArgGlnAlaGlnProProAlaPro.....M 36

438 ACTTCGGCTCCACAAGACCTTCAGCCACAGAAGACCCCTCAATCAGAC 487

36 etAlaAlaArgArgSerGlnArgArgGlyArgArgGlyProGly 52

488 TGTGTCTCGGAAGCTGAGAGCCGACCTGAGACATGATGATACCGAGA 537

53 ThrAlaLeuLeuVal.....ProLeuAlaLeuGlyLeuGlyAlaLe 67

538 CGAGACACCATGTCGAGCCCTCAGACTTCGCTACCGCCGCGCT 587

67 u.....AlaCysLeuGlyLeuLeuAlaValAlaVal 78

588 CCCCCGGCCAGACGTCGAGGTGATCTCCGCTTCAAGAGACTGATG 637

78 erLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94

638 GAGCTGTCGA..... 648

95 LeuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThr 111

649GGCCGCGAGCGAGAGACCG.....GA 671

111 uGlySerGlnAspProAlaProPheLeuAsnArg.....LeuVal 124

672 AGAAGCCGCTATCTTCGCGCACAGCTTCGCGGACATGCTCGCTGAG 721

124 AlArgProArgArgSerAlaProGlySerArgGlyThrArgAlaArg 140

722 TTGCTCCGGAACACTCCGCC.....GGCGTGGCGGCGAGCAGATCA 765

141 AlAlaAlaAlaHisTyrGluValHisProArgProGlyGln..... 154

766 GCG.....CTTCGCTGCTGTCGCGCGCAGCG 791

155AspGlyAlaGln..... 158

792 TCCCCGCGGCTTCCTGAGCCGGTGGCAGACTTCGCGTCCGAGCAG 841

159AlaGlyValAspGlyThrValSer 166

842 ATCTTACGTCGCGAGAGACCGCGCTGGCCAGCGGAGCAGATGGAG 891

167 .GlyTyrGluGluAlaArgGlyLeuAsnSerSerProLeuArgTyrAsn 183

892 GAGCTTCGAGAGCCCATCGAATCTCCGTCGCGCGCTGTCGGGC 941

183 rglInIlleglGluPheLeuValThrArgAlaGlyLeuTyrTyrLeuTyr 199

942 GCTTCGAGGCGCCCTGCTGTCACAGGAGCGGAATCAC..... 982

200 CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 216

983TCCGCGTCCGCGCAGCAGCATGGAGCG 1008

216 uLeuValAspGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAla 233

1009 CTTCCTCGCGCGCTCGGCTCCGCGAGCGCGGAGCCCTTCAGAGAC 1058

233 hrAlaAlaSerSerLeuGly.....ProGlnLeuArgLysCys 245

1059 GGGCGCTCCCAAGATGGAGCTTCGCGGCGCGGATGATGATCATGAC 1108

246 GlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIle 262

1109 TACATCAAGCGGCGTC.....GGCAACAGAGAGCGCGTGGC 1143

262 gThrLeuProThrAlaHis.....LeuLysAlaAlaProPheLeuThrTyr 278

1144 GCTGTGTCTTGGGGCGAAGACTTCGACGCGCGCCGAGATGCGGCGT 1193

278 hegly 279
:::111
1194 ACGCG 1198

seq_name: /SISL/gcdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV69289

seq_documentation_block:
ID: AAV69289 standard; cDNA: 2832 BP.

AAV69289;

01-MAR-1999 (first entry)

Murine JNK-Interacting protein 1 (JIP-1) cDNA.

JIP-1: JNK-Interacting protein; c-Jun NH2-terminal kinase;
signal transduction; inhibitor; mouse; neurodegenerative disease;
Parkinson's disease; Alzheimer's disease; blood clot; stroke;
malignancy; cancer; leukaemia; autoimmune disease; inflammation;
apoptosis; therapy; diagnosis; ds.

Mus sp.

Key Location/Qualifiers
CDS 180..2162
/*tag= a

W09849188-A1.

05-NOV-1998.

28-APR-1998; 98WO-US08513.

28-APR-1997; 97US-0819177.

(UYMA-) UNIV MASSACHUSETTS.

Davis RJ, Dickens M;

WPI: 1999-024042/02.

P-PSDB: AAW81525.

c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to
treat neurodegenerative disease, blood clot, leukaemia, autoimmune
disease, and inflammation

Claim 11: Page 66-70; 95pp; English.

This cDNA sequence codes for murine JNK-interacting protein 1
(JIP-1, see AAW81525), a novel cytoplasmic anchor protein that
specifically binds to and inhibits the biological effects of JNK
(c-Jun NH2-terminal kinase), including the initiation of apoptosis
and oncogenic transformation. No isolate the clone, a yeast
two-hybrid method was used to screen a mouse embryo cDNA library to
identify proteins that interact with JNK. 7 overlapping partial
clones were isolated, and the full-length cDNA was subsequently
obtained by screening a mouse brain lambda ZAPI cDNA library. The
invention provides JIP-1 nucleic acids and polypeptides, expression
vectors and host cells. The JIP-1 polypeptide and nucleic acids
(including antisense and ribozyme molecules) can be used in the
manufacture of a medicament for treating a pathological condition
associated with abnormal expression or activity of JNK, such as a
neurodegenerative disease (selected from Parkinson's disease and
Alzheimer's disease), a blood clot, stroke, malignancy, leukaemia,
an autoimmune disease or inflammation (all claimed).

Sequence 2832 BP; 569 A; 888 C; 840 G; 535 T; 0 other;

alignment_scores:
Quality: 124.50 Length: 263
Ratio: 1.064 Gaps: 12
Percent Similarity: 44.487 Percent Identity: 25.475

alignment_block:
US-09-245-198a-4 x AAV69289

Align seg 1/1 to: AAV69289 from: 1 to: 2832

```

1 MetSerLeuLeuAspPheGluIleSerAlaArgArgLeuProLeuProAr 17
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 ATAACTTGATATCGCTGCGAGACCGCG.....CCGCCCCAGC 60
17 gSerLeuGlySerArgAspGlyAlaValArgGlnAla..... 30
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 TCAGTCCGAACCCGCGCGCGGCGCTCTCCACGCGCTCCGCTGC 110
31 .....GlnProAlaProMetaAlaAlaArg 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
111 TCCCGCGCGCGCGCGCGCGCGCTCCGCGCGCTCCGCGCGGATG 160
41 SerGln.....ArgArgArgGlyArgArgGlyGluProGlyThrAl 54
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 GCCAGGCGTGCACCCCGAATGCGGAGAGAGAGAGAGCGGCTGGCG 210
54 aleuLeuValProLeuAla.....LeuGlyLeu..... 63
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 GCGCGCGCTCCCGCGCGCGCTCCCGCTCCCGCTGAGATGCAATCG 260
63 ..... 63
261 TCCGCTCCCAATTTCAGGCTCACCCATGATCAGCTGAGAGATTGCA 310
64 .....GlyLeuAlaLeuAlaC 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
311 GGATGAAGACCTTTCGAGATCAGTACAGATGTCATCAGCTGAGT 360
69 ys.....LeuGlyLeu.....LeuLeuAlaVal 76
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 GCAAGACACCTGTCTCTCGCGCGCGCGCGCGCGCGCTGTCTGG 410
77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGln 93
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
411 GGTACGAGCGGCGCGCGGAGCGGCTGCAGGCGGAGATCTGCAGAT 460
93 uGluLeuVal.....AlaGluGluAspGln 102
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
461 GGACCTGTCTGCAGCGCGCGAGTGCACATCCGCGCGCGCGAGAG 510
102 sp.....ProSerGluLeuAsnPro 108
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
511 AGGAGAGGAGCAGCAGCTGCTGCCCAAGCAGACAGAGTGGCGCTCC 560
109 GlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValAr 125
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 AAGGCGATCCACACGAGATCCGGGCT.....CG 592
125 gProArgArgSerAlaPro.....LysGlyArgGlyThrArgAla 139
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
593 CAGCGGCGCGCGCGCGCGCGAGCGAGCGAGACCTACCGACCA 642
139 rArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 155
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
643 AGGAGGCTTCCAGCGCTCAACCTTTTCCCGCAGAGTCCGCTCAGAG 692
156 GlyAlaGlnAlaGlyValAspGlyThrValSerGlyThrGluAlaIleAr 172
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
693 ACGCTGAATTAATTAATCTTTAGGCAAAAGACACAGTGGCAGACCGTGT 742
172 gIleAsnSerSerSerProLeuArgTyrAsnArgGln 184
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
743 GTCTCGATCATCTCCCTCTGAAGACAGAGAGACAG 779

```

seq_name: /SISL/gcdata/geneseq/geneseqn-emb1/NA2002.DAT:AAU25519
seq_documentation_block:


```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAA09686
seq_documentation_block:
ID AAA09686 standard; DNA: 3957 BP.
XX
XX AAA09686:
XX
XX 31-JAN-2001 (first entry)
XX
XX HSV-2 immediate early protein ICP4 DNA sequence.
XX
XX Herpes-simplex-virus type 2; HSV-2; Infected cell protein 4; ICP4;
XX KW vaccine; infection; ds.
XX
XX Herpes simplex virus type 2.
XX
XX W09516779-A1.
XX
XX 22-JUN-1995.
XX
XX 13-DEC-1994: 94MO-EP04138.
XX
XX 14-DEC-1993: 93GB-0025496.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Pala P, Cheysen DR, Slaout MM, Koutsoukos MC;
XX
XX WPI: 2001-024142/03.
XX
XX P-PSDB: AAB26874.
XX
XX Immediate early herpes-simplex-virus type 2 (HSV-2) ICP4 protein is
XX PT used in vaccines for therapeutically or prophylactically treating HSV
XX infections -
XX
XX Claim 5; Page 16; 28pp; English.
XX
XX
XX This invention relates to an immediate early herpes-simplex-virus type 2
XX CC (HSV-2) infected cell protein 4 (ICP4) recognised by human cytotoxic T
XX CC cells. HSV-2 ICP4 protein is recognized by cytotoxic T-lymphocyte (CTL)
XX CC cells in humans and is used in vaccines for therapeutically or
XX CC prophylactically treating HSV infections. Pharmaceutical compositions of
XX CC HSV-2 ICP4 protein may be used to treat patients suffering from HSV
XX CC infections, to prevent or decrease recurrent herpes disease, frequency,
XX severity and duration of episodes. The present sequence represents HSV-2
XX DNA encoding ICP4.
XX
XX
XX Sequence 3957 BP; 368 A; 1656 C; 1568 G; 365 T; 0 other;
XX
XX
XX Alignment_scores:
XX Quality: 124.00 Length: 171
XX Ratio: 1.610 Gaps: 8
XX Percent Similarity: 45.029 Percent Identity: 30.409
XX
XX
XX alignment_block:
XX US-09-245-198A x AAA09686 ..
XX
XX
XX Align seg 1/1 to: AAA09686 from: 1 to: 3957
XX
XX 10 AATATGATGLeuProLeuProArgSerLeuGlySerArgAspGlyGlyAl 26
XX ||||||| ||| ||| :||| ||| ||| ||
XX GCGCCGCGCGCGCGCGCGCGCGCGCGCGCGG...CGCGCGCGCGCGC 2053
XX
XX 26 AATATGATGAlaGlnProProAlaProMetAlaAlaArg..... 39
XX ||:||||: ||||||| :||| |||||
XX GCGACGCGACGCGCGCGCGCGCGCGCGCTGCGTCTGCGGACGTGCGGTTTC 2103
XX
XX 40 .....ArgSer 41

```

```

2104 GTGGCCGACGCGCTGTGTCATGCAGTCCGCCGCGGGAGACTCGCGCGTGCC   |||:::
      42 GlnArgArgArgGlyArgArgGly.....GluProGlyTh 53
      2154 CGCGCGCAGCGAGCGCGCGCTGTGCGCCGCGCGCCGTGAGCGCTGGTGC 2203
          53 rAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysL 70
      2204 CCGG.....GGCCCTGGGCGCGC..... 2222
          70 euGLyLeuLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86
      2223 .....GCT 2225
          87 AlaGlnGluProAlaGlnGlnGluLeuValAlaGlnGluAspGlnAspr 103
      2226 GCCCGGAGCGCGCGCGCTGTGAGCTCCGCCGCGCGCGCGCGGAGCC 2275
          103 oSerGluLeuAsnProGlnThrGlnGluSerGlnAspProAlaProPheL 120
      2276 TGCTTTCTCCAGACCA.....GAGCCTCGCGCCCCCTG 2307
          120 euASnArgLeuValArgProArgArgSerAlaProLySGlyArgLySthr 126
      2308 CTGGC.....CGACACCCTGCGCGCGCGCGAGCTGCTGCGCGGCCC 2348
          137 ArgAlaArgArgAlaIleAlaIleHisTyrgLyuValHisProArgProG 153
      2349 CGCCTCGCGCGCGCGCGAGCGCGCCAGAAGCGA...GAGCGCGCGCGCG 2395
          153 yGIAspGlyAla 157
              |||| ::|||
              2396 CCAGCGCGCGCGCC 2408

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAD25519

seq_documentation_block:
ID_AAD25519 standard; DNA; 154746 BP.
XX AC AAD25519;
XX DT 26-MAR-2002 (first entry)
XX DE Human herpesvirus 2 complete DNA genome.
XX KW Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;
    antibacterial; fungicidal; protozoacidal; antirheumatic; antiinflammatory;
    antidiabetic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
    immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
    vasculitis; ds.
XX OS Human herpesvirus 2.
XX PN WO200176643-A1.
XX PD 18-OCT-2001.
XX PE 06-APR-2001; 2001WO-US11372.
XX PR 07-APR-2000; 2000US-195860P.
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PI Orson FM, Kinsey BM, Bhogal BS;
XX WPI: 2002-066308/09.
XX DR
XX PT Composition for oral delivery of vaccines, comprises expression vector
    containing antigenic genomic sequence, bound to aggregated
    protein-polycationic polymer conjugate or suspension -
XX PS Disclosure; Page 90-132; 145pp; English.
```


THIS PAGE BLANK (USPTO)

OM of: US-09-245-198a-4 to: GenEmbl:* out-format: pfs
Date: Jun 22, 2002 2:30 PM
About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frim+ p2n.model -DEV=xlh
-O/cgml2.1/USPTO.spool/US09245198/rnat.21062002.184006.9945/app_query.fasta.1.632
-DB=GenEmbl -QFMT=fastap -SUFFIX=trg -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DEEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09245198 -CGML1.1.0697
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:
Query: US-09-245-198a-4
Query length: 284
Database: GenEmbl.*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 2600.680000

score list:	Strid	Orig	zscore	EScore	len	Documentation
gb_pat:AX201324	+ 1366.00	1243.78	6.5e-61	1353	AX201324	Sequence 3 from Patent
gb_pr:AF055872	+ 1366.00	1243.69	6.5e-61	1368	AF055872	Homo sapiens Apol/D3
gb_pr:BC019047	+ 1331.50	1211.08	4.3e-59	1651	BC019047	Homo sapiens, Similar
gb_pr:AF030099	+ 1298.00	1182.65	1.6e-57	1306	AF030099	Homo sapiens, TWEAK mR
gb_pat:ARL40407	+ 1265.00	1153.27	7.1e-56	1236	ARL40407	Sequence 1 from Patent
gb_pat:AX180714	+ 1066.00	976.26	5.3e-46	898	AX180714	Sequence 1 from Patent
gb_ro:AF030100	+ 1020.00	932.48	1.4e-43	1168	AF030100	Mus musculus TWEAK mR
gb_hlg:AC016876	- 589.50	504.24	1.0e-19	19038	AC016876	Homo sapiens c10ne
gb_hlg:AC069459	- 555.50	473.03	5.5e-18	203083	AC069459	Mus musculus c10ne
gb_ro:AL603707	- 555.50	471.93	6.3e-18	234182	AL603707	Mouse DNA sequence
gb_hlg:AC098923	- 385.50	320.50	1.7e-09	179030	AC098923	Rattus norvegicus c
gb_pl:AF002482	+ 145.50	103.41	2.1e+03	187835	AF002482	Oryza sativa genome
gb_ba:SC1G2	- 140.00	103.03	912.82	42210	AL031124	Streptomyces coelic
gb_ba:AL646059	- 138.50	97.00	4.9e+03	190050	AL646059	Rattus norvegicus c
gb_hlg:AC105470	+ 136.50	93.59	7.5e+03	233912	AC105470	Rattus norvegicus c
gb_ba:AF002222	+ 135.00	120.15	249.27	6397	AF002222	Pseudomonas aerugin
gb_ba:AE004651	+ 135.00	116.14	416.92	10728	AE004651	Pseudomonas aerugin
gb_ba:AE004740	- 134.00	115.06	478.40	10965	AE004740	Pseudomonas aerugin
gb_ba:SCC61A	- 130.00	101.35	2.8e+03	40352	AL356595	Streptomyces coelic
gb_ba:SCDMSSTS	+ 129.00	115.56	449.22	5751	X89010	S. glaucus genes str
gb_ba:SCC46	+ 129.00	107.53	1.3e+03	16185	AL139164	Streptomyces coelic
gb_ba:SCG10	- 129.00	104.02	2.0e+03	25459	AL006985	Streptomyces glauc
gb_ba:SCG10	- 129.00	102.02	2.5e+03	32941	AL133469	Streptomyces coelic
gb_pr:CMS01DV6	+ 129.00	99.78	3.4e+03	43961	AL133422	Streptomyces coelic
gb_pr:AC016708	- 127.00	86.45	1.9e+04	194304	AC016708	Homo sapiens BAC c1
gb_ba:AL646063	+ 126.50	85.50	2.1e+04	207050	AL646063	Rattus norvegicus c
gb_pr:HSBARD11	+ 125.50	85.97	118.10	999	AF038034	Homo sapiens BRCAL-ase
gb_ba:AVIN1FC	- 125.50	99.90	3.3e+03	28795	M20568	A. vinelandii mfor nlf
gb_hlg:AP003822	- 125.50	91.31	1.0e+04	87143	AF003822	Oryza sativa chromo
gb_hlg:AP004272	+ 125.50	86.91	1.8e+04	153555	AP004272	Oryza sativa chromo
gb_ba:SCG42	+ 125.00	98.98	3.8e+03	30590	AL031371	Streptomyces coelic
gb_ba:SCG610	- 124.50	97.11	4.8e+03	36734	AL049497	Streptomyces coelic
gb_hlg:AP0033876	- 124.50	89.46	1.3e+04	98504	AP003876	Oryza sativa chromo
gb_hlg:AP003911	- 124.50	88.81	1.4e+04	107132	AP003911	Oryza sativa chromo
gb_v1:HSV2HG52	- 124.50	85.95	2.0e+04	154746	286099	Herpes simplex virus
gb_hlg:AP003634	+ 124.00	113.94	3.3e+04	176580	AP003634	Oryza sativa chromo
gb_pat:AA5258	+ 124.00	113.94	552.59	3957	AA5258	Sequence 2 from Patent
gb_pr:HS745C22	- 124.00	104.17	1.9e+03	13949	AL031596	Human DNA sequence
gb_v1:HSV2HG52	+ 124.00	85.50	2.1e+04	154746	286099	Herpes simplex virus

gb_ba:AE004731	+ 123.50	101.42	2.8e+03	18765	AE004731	Pseudomonas aerug
gb_hlg:AP003979	- 123.50	85.97	2.0e+04	137366	AP003979	Oryza sativa chr
gb_ro:MAAF003115	+ 123.00	115.75	438.25	2790	AF003115	Mus musculus JIP-1
gb_ro:AF054611	+ 123.00	115.37	460.28	2931	AF054611	Mus musculus JNK 1
gb_ba:AB070956	- 122.50	92.30	8.9e+03	54101	AB070956	Streptomyces aver
seq_name: gb_pat:AX201324						
seq documentation block:						
LOCUS	AX201324	1353 bp	DNA	linear	PAT 30-AUG-2001	
DEFINITION	Sequence 3 from Patent WO0153486.					
ACCESSION	AX201324					
VERSION	AX201324.1	GI:15391154				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1353)					
AUTHORS	Emmalyola, Metaxoa, Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthelia; Primates; Catarrhini; Homnidae; Homo. Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J., Pitti, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I. Compositions and methods for the treatment of tumour Patent: WO 0153486-A 3 26-JUL-2001; Genentech, Inc. (US)					
TITLE	Location/Qualifiers					
JOURNAL	1. 1353					
FEATURES	/organism="Homo sapiens" /db_xref="taxon:9606"					
SOURCE	257 a 443 c 389 g 264 t					
BASE COUNT						
ORIGIN						
alignment scores:						
Quality:	1366.00	Length:	268			
Ratio:	5.097	Gaps:	0			
Percent Similarity:	100.000	Percent Identity:	100.000			
alignment block:						
US-09-245-198a-4 x AX201324						
Align seg 1/1 to: AX201324 from: 1 to: 1353						
17 ArgSerLeuGlySerArgaspGlyGlyAlaValArgGlnAlaGlnProPr 33						
1 CGATCCCTGGGTCCTCCGGATGGGGGGCGGTGAGGACAGACAGCCCC 50						
33 oAlaProMetaAlaAlaArgArgSerGlnArgArgArgGlyArgGlyG 50						
51 CGCCCGCAGCGCCGCCCGCGCGAGCCAGAGCGGAGGCGCGCGGGGG 100						
50 InProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuAla 66						
101 ACCCGGGACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 150						
67 LeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArg 83						
151 CTGGCCCTGGCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200						
83 aSerLeuSerAlaGlnGluProAlaGlnGluLeuValAlaGlnGlu 100						
201 ATCGCGTCGCGCCAGAGAGCGCGCCAGAGAGCGCGCGCGCGCG 250						
100 spGlnAspProSerGluLeuAsnProGlnThrGluGlnSerGlnAsp 116						
251 ACCAGAGACCGCGCGAGAAATCCCGACAGAGAAAGCCAGATCCT 300						
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLys 133						
301 GCGCCCTTCTTACCGACTAGTTGGCTCCGAGAGAGCGACCTTAAG 350						
133 yArgLysThrArgAlaArgArgAlaIleAlaIleHisLysGluValHis 150						

```
351 CCGGAAAACGCGGCTCGAAGAGCGATCGACCCCATTTATGAAGTTCATC 400
150 roatpTroglyGlnaSpGlyAlaGlnaIaGlyValaSpGlyThrValSer 166
|||||
401 CACGACTGACAGACGAGCGGAGCGGAGCTGTGGACGGGACAGTACT 450
167 GLYTPGlnGlnaAlaArgIleasnSerSerProleuArgTyrAsnAr 183
|||||
451 GCGTGGAGAGCGCAGATCAACAGCTCCAGCCCTCGCGCTACACACG 500
183 gGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyr 200
501 CCAAGATCGGGAGATTATAGTACCCCGGCTGGCTTACTACTACTACT 550
200 ysgInValHisPheaspGlyGlyAlaValTyrLeuLysLeuAspLeu 216
|||||
551 GTGAGGTGCACTTTGATGAGGAGAGCTGTACTCTGAAGCTGACTTG 600
217 LeuValaSpGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaTh 233
|||||
601 CTGGTGGATGTGTGTGGCCCTGCTGCTGAGGAATTCACAGCCAC 650
233 rAlaAlaSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
|||||
651 TCGGGCGAGTTCCTCGGCGCCCAAGCTCCCTCTGCGAGGTGTCTGGC 700
250 euleuAlaAlaArgProGlySerSerLeuArgIleArgThrLeuProTyr 266
|||||
701 TGTGGCGCCCGCGGCGGAGGCTCCCTGCGGAGTCCGACCCCTCCCTGG 750
267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
|||||
751 GCCCATCTCAAGGCTGCCCTTCCTCCTCCTACTCGACTCTTCAGGT 800
283 His 284
|||||
801 TCAC 804
```

seq_name: gb-pr:AF055872

seq_documentation_block:

LOCUS AF055872 1368 bp mRNA linear PRI 04-MAY-1998

DEFINITION Homo sapiens ApO3/DR3 ligand (APO3L) mRNA, complete cds.

ACCESSION AF055872

VERSION AF055872.1 GI:3108230

KEYWORDS

SOURCE

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1368)

Marsters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and

Ashkenazi,A. Identification of a ligand for the death-domain-containing receptor

ApO3 Curr. Biol. 8 (9), 525-528 (1998)

JOURNAL MEDLINE 98228355

REFERENCE 2 (bases 1 to 1368)

Marsters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and

Ashkenazi,A. Direct Submission

Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,

South San Francisco, CA 94080, USA

FEATURES

source 1..1368

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17p13"

gene 1..1368

/gene="APO3L"

CDS 58..807

/gene="APO3L"

```
/function="binds to the death-domain-containing receptor
ApO3/DR3"
/note="TNF homolog"
/codon_start=1
/product="ApO3/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARSORRRGRGRRGPGTALLPLATLGLALGILLAVSL
GSRSLAEOEPAOEELVAEEDODPSLELPTOEESODPELNLVPRPRSPKGRKTR
ARRAIAHVEVHPRGODCAQAGVTVSGSEFARINSSPLRYNROGERTVRAGL
YLLQVHFDESKAVYLLKLDLVDSVALRLEFSRATMASSLGQRLRCOVSGLLAL
RPGSLRLRLTPMAHLKAAPFLYFGLEQVH"
BASE COUNT 272 a 443 c 389 g 264 t
ORIGIN
```

```
alignment_scores:
Quality: 1366.00 Length: 268
Ratio: 5.097 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-245-198a-4 x AF055872 ..

```
Align seg 1/1 to: AF055872 from: 1 to: 1368
17 ArgSerLeuGlySerArgAspGlyGlyAlaValArgGlnaIaGlnProPr 33
1 CGATCCCTCGGATCCCGGATGGGGGGCGGTGAGCAGGCACAGCCCC 50
33 oAlaProMetaIaAlaArgArgSerGlnArgArgArgGlyArgArgGly 50
51 CGCCCCATGGCGCCCGCTCGAGACAGAGCGAGGGGGCGCGGGGGGG 100
50 luProGlyThrAlaLeuLeuValProleuAlaLeuGlyLeuGlyLeuAla 66
101 AGCCGGGACCGCCCTGCTGCTCCCGCTGCGCTGGCGCTGGCGCTGGCG 150
67 LeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAl 83
151 CTGGGCTGCTCGGCTCCCTGCTGCGCGGTGCTGAGTTTGGGGAGCCGGC 200
83 aSerLeuSerAlaGlnGluProAlaGlnGlnGluLeuValaIaGlnGlu 100
201 ATGCGCTGCCGCGCAGAGAGCTGCCAGAGAGTGTGTCAGAGAGAGG 250
100 spGlnaSerProSerGlnLeuAsnProGlnThrGlnGlnSerGlnaPro 116
251 ACCAGGACCGGTGGAACTGAATCCCGACAGAAAGAACCGAGATCTT 300
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysG 133
301 GCGGCTTTCCTGAACGACAGTACGTCGCGCTCGCAGAAAGTCCACTTAAG 350
133 yArgLysThrArgAlaArgArgAlaIleAlaIleHisTyrGlyValHisP 150
351 CCGGAAAACAGGGCTCGAAGAGCGATCCGACGCCCATTTATGAAGTTCATC 400
150 roatpTroglyGlnaSpGlyAlaGlnaIaGlyValaSpGlyThrValSer 166
401 CACGACTGACAGACGAGCGGAGCGGAGCTGTGGACGGGACAGTACTG 450
167 GLYTPGlnGlnaAlaArgIleasnSerSerProleuArgTyrAsnAr 183
|||||
451 GCGTGGAGAGCGCAGATCAACAGCTCCAGCCCTCGCGCTACACACG 500
183 gGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyr 200
501 CCAAGATCGGGAGATTATAGTACCCCGGCTGGCTTACTACTACTACT 550
200 ysgInValHisPheaspGlyGlyAlaValTyrLeuLysLeuAspLeu 216
|||||
551 GTGAGGTGCACTTTGATGAGGAGAGCTGTACTCTGAGCTGACTTG 600
```


217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThr 233
|||||
601 CTGGTGAATGCTGTCTGGCCCTGGCCCTGCTGAGAAATTCACACCAC 650
233 rAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
|||||
651 TCGGCGCAATCTCCCTCGGGCCCGCAGCTCCGCTCTCCAGAGTCTGGGC 700
250 euLeuAlaLeuArgProGlySerSerLeuArgGlnLeuArgThrLeuProTrp 266
|||||
701 TGTGGCCCTGGCGCCAGGCTCTCCCTCGGATCCGACCCCTCCCTGG 750
267 AlaHisLeuValAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
|||||
751 GCCCATCTCAAGGCTGCCCCCTTCTCACCCTACTTGGACTTTCAGGT 800
283 His 284
|||||
801 TCAC 804
seq_name: gb_pr:BC019047
seq_documentation_block:
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, Similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047.1 GI:17512138
VERSION BC019047.1
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-T., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C.,
McOmwell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, D.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 30 Row: P Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507596.
Location/Qualifiers
1..1651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20669 IMAGE:4766071"
/tissue_type="Primary B-Cells from Tonsils"
/clone_lib="NIH_MGC_48"

/lab_host="DH10B-R"
/note="Vector: pOTB7"
106..510
CDS
/codon_start=1
/product="Similar to tumor necrosis factor (ligand)
superfamily, member 12"
/protein_id="AAH19047.1"
/db_xref="GI:17512139"
/translation="MAARRSRRRGRGREGPTALLVPLALGLALACGLLAVSL
GSRASLSAEPDAEELVAEDQDPSLNPOTSESQDPAPPLNLRVPRRSAPGKTR
ARRAIAAHYEHHPRGDGAADADGGYTCARP"
BASE COUNT 344 a 517 c 481 g 309 t
ORIGIN
alignment_scores:
Quality: 1331.50 Length: 356
Ratio: 4.84 Gaps: 2
Percent Similarity: 77.247 Percent Identity: 77.247
alignment_block:
US-09-245-198a-4 x BC019047 ..
Align seg 1/1 to: BC019047 from: 1 to: 1651
9 SerAlaArgLeuProLeuProArgSerLeuGlySerArgAspGlyL 25
|||||
25 TCGGCCCGCGGCTCCCTCCCTCCCGATCCCTGGGCTCCGGGATCGGG 74
25 yAlaValArgGlnAlaGlnProProAlaProMetaAlaAlaArgSerG 42
|||||
75 GCGGTGAGAGGACGACACGCCCGCCCGCATGGCGCGCCGTGGAGGC 124
42 InArgArgArgGlyArgArgGlyGlyProGlyThrAlaLeuValPro 58
|||||
125 AGAGCGGAGGAGGCGCCCGGGGGAGCCGGGACCCGCTGCTGCCG 174
59 LeuAlaLeuGlyLeuGlyLeuAlaLeuAlaGlySerGlyLeuLeuVal 75
|||||
175 CTGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 224
75 aValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaG 92
|||||
225 CGGTGCTCAATTTGGGGAGCGCGGCATCGCTGCCCGCAGAGCGTCCC 274
92 InGluGluLeuValAlaGluGluAspGlnAspProSerGluLeuAsnPro 108
|||||
275 AGGAGGAGCTGGGCGCAGAGAGACAGACCGCTCGGAATGATCCC 324
109 GlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuVal 125
|||||
325 CAGACGAGAAAGAACCGAGATCCCTGCTTCCTGACCGATGATTCG 374
125 gProArgArgSerAlaProLeuGlyArgGlyThrArgAlaArgArgAla 142
|||||
375 GCGTGCACAAGTGCACCTRAAGGCGGAAACACGCGCTCGAAGCGCA 424
142 LeuAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGln 158
|||||
425 TCGCAGCCCATATGAAATTCATCCAGACCTGACAGACGAGAGCGAG 474
159 ..Ala..... 159
475 GCAGATGAGAGTTACACAACCTGTCTGAGGCCATGAGACTAAGTGTCG 524
159 159
525 GAGCCAAAGATTGAACCAAGTAGAATGTGCTGTACTGACATGTCG 574
159 159
575 TGCATGAAGAGGAGGTGCTGTGTCAGAGGGTGAGGGGTCCATGACAGGG 624

159 159
625 CCACATCCAAAAAGGAGAGAGAAAGTTTCCAGAAAGAGAGAGACACA 674
160GlyValAs 162
675 TCTCCACACATTACCAGAGGTCAAAGAGACAAAGACCAGGAGTGGGA 724
162 pGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSerProL 179
725 CGGACAGTGAAGTGGCTGGAGGAGGAGGAGATCAACAGCTCCAGCCCTC 774
179 euArgTyraSerArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
775 TGCCTACACACCCAGATCGGGAGATTATAGTCACCCGGGCTGGGCTC 824
196 TyTrpLeuTyrcysGlnValHisPheAspGluGlyLysAlaValTyrLe 212
825 TACTACTGTAAGTCAAGGTCACCTTGTATGAGGAGGAGGAGGCTGTA 874
212 uLysLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGlu 229
875 GAACCTGGACTGCTGGTGGATGGTGGCTGGCCCTGGCTGGAGG 924
229 IupheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeucys 245
925 AATTCTCAGCCATCGCGGAGTTCCCTCGGGCCCGCCAGCTCGGCTGTC 974
246 GlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleAr 262
975 CAGGTGCTGGGCTGTGGCCCTCGGGCAGGCTCCTCGCGGAGTCG 1024
262 gThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPhe 279
1025 CACCTCCCGGGGCGCATCTCAAGGCTGCCCTTCCTCCTCAGCTTGC 1074
279 IyLeuPheGlnValHis 284
1075 GACTCTTCCAGGTTCC 1091
seq_name: gb_pr:AF030099
seq_documentation_block: 1306 bp mRNA linear PRI 20-DEC-1997
LOCUS AF030099
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1306)
TITLE Chicheportiche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,
Hession, C., Garcia, I. and Browning, J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J. Biol. Chem. 272 (51), 32401-32410 (1997)
98070415
2 (bases 1 to 1306)
Bourdon, P., Hession, C., Tizard, R. and Browning, J.
Direct Submission
Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
source
1..1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/feature_type="tss11"
18..767
/note="Ligand in the TNF family; secreted protein; start

codon not verified experimentally"
/codon_start=1
/product="TWEAK"
/protein_id="AAC51923.1"
/db_xref="GI:2707218"
/translation="MAARSSQRRRRRREPEPTALLVPLALGLALACTGLLAVSL
GSRAVLSAQEPAGEELVAEEDDPSELNPQTEESDPAFLNRLVRRPRSPKGRTR
ARRAIAHYEVHPRGODGAQAGVDGTVSGWEARI NSSPLRVNRQIGETIVRAGL
YLYCOVHEDEKAVYLRKLDLVDVGLARCLIEERSATVAASLIGPOLRLCOVSLAL
RPGSLRIRTRTPMAHLKAPFLTYRGLPOVH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN
alignment_scores:
Quality: 1298.00 Length: 254
Ratio: 5.110 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-245-198a-4 x AF030099 ..
Align seg 1/1 to: AF030099 from: 1 to: 1306
31 GlnProProAlaProMetAlaAlaArgArgSerGlnArgArgGlyArg 47
3 CAGCCCCCCCCCATGGCCGCGCCGCGAGCCAGAGCGGAGGCGGCG 52
47 gArgGlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeu 64
53 CCGGGGGAGCGGGGACCGCCCTGCTGTCGCTCGCTGCGCTGGGCTGG 102
64 IyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGly 80
103 GCCTGGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGG 152
81 SerArgAlaSerLeuSerArgAlaGlnGluProAlaGlnGluLeuValAl 97
153 AGCCGGGATCGCTGCTCGCCGCGAGGAGCTGCCAGAGAGAGCTGGGC 202
97 gLugLusAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerG 114
203 AGAGAGAGACAGGAGCCGTCGGAATGAAATCCCAAGACAGAAAGACC 252
114 IAspProAlaProPheLeuAsnArgLeuValArgProAlaArgSerAla 130
253 AGATCTCGCGCTTCTCTGTAACGCACTAGTTCGCTGCCAGAAAGTCA 302
131 ProLysGlyArgGlyThrArgAlaArgAlaIleAlaHisTyrGln 147
303 CCTAAAGCGCGAAACACGGGCTCGAAAGCATCCAGCCCATTTATGA 352
147 uValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspTyr 164
353 AGTTCAATCCAGACTGTGACAGCGAGGCGCAGGCTGAGTGGACGGGA 402
164 hrValSerGlyTrpGluGluAlaArgIleAsnSerSerProLeuArg 180
403 CAGTGAAGTGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
181 TyraSerArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 197
453 TACAACCGCCAGATCGGGAGATTATAGTCAACCGGCTGGGCTGTA 502
197 rLeuTyrcysGlnValHisPheAspGluGlyValAlaValTyrLeuLys 214
503 CCTTACTGTCAGGTGCACTTGTATGAGGAGGAGGAGGAGGAGGAGGAG 552
214 euAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPhe 230
553 TGGACTTGGCTGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
231 SerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeucysGlnVa 247

```
|||||
603 TCAGCCACTGCGCCAGTTCCTCGGGCCCGACCTCGGCTTGCCAGGT 652
247 |SergIyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrL 264
|||||
653 GTCCTGGGCTTTGGCCCTCGGGCAGGATCTCCGCGATCGCAGCC 702
264 eupProTpaHhIsLeuLysAlaAlaProPheLeuThrThyPhgIyLeu 280
|||||
703 TCCCTCGGGCCCACTCTCAAGGCTGCCCTTCTCTCACTTCTGAGCTC 752
281 phgIuValHis 284
|||||
753 TTCAGAGTTCAC 764

seq_name: gb_pat:AR140407
seq_documentation_block:
LOCUS AR140407 1236 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207642.
ACCESSION AR140407
VERSION AR140407.1 GI:14482903
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1236)
AUTHORS
TITLE Member of the TNF family useful for treatment and diagnosis of
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
SOURCE 1. 1236
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN

alignment_scores:
Quality: 1265.00 Length: 249
Ratio: 5.080 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.598

alignment_block:
US-09-245-198a-4 x AR140407 ..
Align seg 1/1 to: AR140407 from: 1 to: 1236
36 MetAlaAlaArgArgSerGlnArgArgGlyArgArgGlyArgGlyArgProG 52
|||||
1 ATGGCGCGCGCGTCGGAGCCAGAAAGCGGAGGGCGCGCGGGGAGCCGG 50
52 yThraAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAla 69
|||||
51 CACCGCGCTGCTGTCGCCGCTCGCGCTGGGCTCGGGCGCTGCGCTG 100
69 yLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeu 85
|||||
101 GCGTCGGCCCTCGCTGCGCTGCTGCTAGTTGGGAGCCGGGCGATCGCT 150
86 SerAlaGlnGluProAlaGlnGluGluValAlaGlnGluAspGlnAs 102
|||||
151 TCCGCGCCAGAGAGCTGCCAGAGAGCTGCTGGCAGAGAGAGAGCCAGGA 200
102 pProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaPro 119
|||||
201 CCCGTCGGAACCTGAATCCCCAGACAGAAAGCCAGGATCTCGCCCTT 250
119 heLeuAsnArgLeuValArgProArgSerAlaProGlyArgGly 135
|||||
251 TCCCTGAACGAGCTAGTTGGCCCTCGAAGAGTGCACCTAAAGCCGGA 300
136 ThrArgAlaArgArgAlaIleAlaIleAlaHisIstYrGluValHisProArgPr 152
```

```
|||||
301 ACACGGGCTCGAAGAGCGATTCGAGCCCAATTATGAGTTCATCAGAC 350
152 oGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSergIyTrp 169
|||||
351 TGGACAGAGACGAGCCAGAGGTGTGAGCAGGACAGTACAGTGGCTGG 400
169 InGluAlaArgIleAsnSerSerProLeuArgIyrIstnArgIle 185
|||||
401 AGGAAGCCAGAAATCAACAGCTCCAGCCCTCTGGCTACAAACGCCAGATC 450
186 GlyGluPheIleValThrArgAlaGlyLeuTyTYrLeuTyTCysGln 202
|||||
451 GGGAGTTTATATGACACCGGGCTGGGCTCTACTACTGCTAGCTAGGT 500
202 HisPheAspGluGlyLysAlaValIyrLeuLysLeuAspLeuVal 219
|||||
501 GCACTTTGATGAGGAGGAGGAGCTGTCTACCTGAAGCTGAGACTTGCT 550
219 spGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAla 235
|||||
551 ATGCTGTGCTGCCCTCGCTGCTGAGGAATTCACACCACTGCGGCG 600
236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSergIyLeuVal 252
|||||
601 AGTTCCTCGGGGCCAGAGCTCCGCCCTCTGCAGGTCTGCGGTGGC 650
252 AlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTpaHhIsL 269
|||||
651 CCTGGCGCCAGAGGCTCTCCCTCGGATCCGACCCCTCGGGCCCATC 700
269 eulysAlaAlaProPheLeuThrThyPhgIyLeuPheGlnValHis 284
|||||
701 TCAAGGCTGCCCCCTTCCTCACTTCTGAGCTTTCCAGGTTCAC 747

seq_name: gb_pat:AX180714
seq_documentation_block:
LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent W00145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 898)
AUTHORS
TITLE Weak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES
SOURCE 1. 898
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=1
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MARGSRSTLLAFGLICTPLDQSGATSSDRMKQIDKITEITS
KIYIENIARIKKLIGERTSSLSGRASVSAQRPQDELYAEEDQPSLEINPOTERS
QDPAPFLNRILVRPRRSAPKGRKTRARAIAAHVHRPPQDDGAQAGVDVSGWEA
RINSSPLRVNRQIGERTVYRAGLYLVYCOVHPDEGRAVYLKLDLVDGVIALCLDEE
FSARAASLSIGPOLRLCOVSGILARLPSSLRIRLTPMAHKAPFLTYFGLFQVH"

BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

alignment_scores:
Quality: 1066.00 Length: 207
Ratio: 5.150 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-4 x AX180714 ..

Align seg 1/1 to: AX180714 from: 1 to: 898

```

78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluG1 94
|||||
250 AGTTGGGAGGCGCGGCGCTGCTCCGCCAGAGGCGTGGCCAGAGGA 299
94 ULLeuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrG 111
|||||
300 GCTGTGGCAGAGGAGGAGCAGGACCGCTGGAACTGAATCCCGAGACAG 349
111 LuGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArg 127
|||||
350 AAGAAACCCAGAGATCTCGCGCTTCTCGAACCCAGCTAGTTGGCTCGC 399
128 ArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAlaAl 144
|||||
400 AGAAGTCACCTAAAGCCGGAACACGGGCTCGAAGAGGATCGCAGC 449
144 AhISTyGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 161
|||||
450 CCATTATGAAGTTTCATCCAGCAGCTGGACAGAGGAGCGAGCAGCTG 499
161 AlaSPGlyThrValSerGlyTyrPGLuGluAlaArgIleAsnSerSer 177
|||||
500 TGGACGGGACAGTAGTGCTGGGAGGAGACCAATCAACAGCTCCAGC 549
178 ProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAla 194
|||||
550 CCTCTGGCTACACACCGCAGATCGGGAGATTATAGTCAACCGCGCTGG 599
194 yLeuTyTYrLeuTyrcysGlnValHisPheAspGluGlyLysAlaVal 211
|||||
600 GCTTACTACTGTGATCTGTCAGGTGACTTTGATGAGGGGAGGCTGTCT 649
211 yrlLeuLysLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeu 227
|||||
650 ACCTGAAGCTGAGTCTGCTGGTGGATGCTGTGCTGGCTGGCTGGCTG 699
228 GLuGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLe 244
|||||
700 GAGGAATTCTCAGCCACTGGCGCAGTTCCCTCGGGCCCCAGCTCCGCT 749
244 uCySGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 261
|||||
750 CTGCCAGGTGCTGTGGCTGTGGCCCTGGCGGCGAGGCTCCCTCGCGGA 799
261 lEarTYrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTy 277
|||||
800 TCCGCACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTAC 849
278 PheGlyLeuPheGlnValHis 284
|||||
850 TTCCGACTCTTCAGAGTTTCAAC 870

```

seq_name: gb_ro:AF030100

seq_documentation_block:

LOCUS AF030100 1168 bp mRNA linear ROD 20-DEC-1997
 DEFINITION Mus musculus TWEAK mRNA, partial cds.
 ACCESSION AF030100
 VERSION AF030100.1 GI:27072220

KEYWORDS

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1168)
 AUTHORS Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,

TITLE Hession,C., Garcia,I. and Browning,J.L.
 JOURNAL TWEAK, a new secreted ligand in the tumor necrosis factor family
 MEDLINE J. Biol. Chem. 272 (51), 32401-32410 (1997)
 REFERENCE 98070415

AUTHORS 2 (bases 1 to 1168)
 TITLE Chicheportliche,Y., Bixler,S., Tizard,R. and Browning,J.
 JOURNAL Direct Submission
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
 Cambridge, MA 02142, USA

FEATURES Location/Qualifiers

source

1..1168
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="peritoneal macrophages"
 <1..679
 /note="ligand in the TNF family; secreted protein"
 /codon_start=2
 /product="TWEAK"
 /protein_id="AAC53517.1"
 /translation="VLSGLALACILLLVYVLSGMATLSAOPSEBELFAEDREP
 PELNPTQESDDVPEFLQIVPRRSAPKGRARPAIAAHYEVHPRPDGAQVY
 DCFVSGMEETKINSSPLRYDROIEFTYIRAGLYLYLCVHPDEGRKAVYIAKLDELIV
 GVLAIRCLSEFSATVASSPQRLRQVSGLLPLRPSSLRIKTLPMANHLAAPELTY
 FELQVH"

CDS

BASE COUNT 242 a 360 c 298 g 268 t
 ORIGIN

alignment_scores:
 Quality: 1020.00 Length: 224
 Ratio: 4.744 Gaps: 0
 Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198a-4 x AF030100 ..

Align seg 1/1 to: AF030100 from: 1 to: 1168

```

61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValVa 77
|||||
5 CTGAGCCTGGGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 54
77 lSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94
|||||
55 CAGCTGGGAGCGTGGGCAAGCGTGTCTGCCAGAGACCTTTCAGAGG 104
94 LuLeuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThr 110
|||||
105 AGCTGACAGCAGAGAGCCGCCGGAGCCCTGTGAATCGAATCCCGAGACA 154
111 GluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAr 127
|||||
155 GAGGAAGCCAGGATGTGTACTTCTTGGAACAACCTAGTCGGCTCG 204
127 gArSerAlaProLysGlyArgLysThrArgAlaArgArgAlaIleAla 144
|||||
205 AAGAAAGTCTCTAAAGCCGGAAGCGCGGCGCTCGCGCAGATTCAG 254
144 lAhISTyGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
|||||
255 CCCATTATGAGGTTCAATCCCGCAGAGAGATGAGTGAAGCAGCAGGT 304
161 ValAspGlyThrValSerGlyTyrPGLuGluAlaArgIleAsnSerSer 177
|||||
305 GTGGATGGACAGTGAAGTGGCTGGGAAGAGACCAAAATCAACAGCTCAG 354
177 rProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAla 194
|||||
355 CCGCTCTGCGTACAGACCGCAGATGGGGAATTTACAGTCAATCAGGCTG 404
194 yLeuTyTYrLeuTyrcysGlnValHisPheAspGluGlyLysAlaVal 210

```



```

misc_feature      /note="assembly_fragment"
68265..94625
misc_feature      /note="assembly_fragment"
94726..102564
misc_feature      /note="assembly_fragment"
102665..116764
misc_feature      /note="assembly_fragment"
116865..136408
misc_feature      /note="assembly_fragment"
136509..156380
misc_feature      /note="assembly_fragment"
156481..180631
misc_feature      /note="assembly_fragment"
180732..190358
misc_feature      /note="assembly_fragment"
clone_end:R7
vector_side:right"

BASE COUNT      45173 a 48605 c 50635 g 44441 t 1504 others
ORIGIN

Alignment_scores:
Quality: 589.50      Length: 192
Ratio: 4.679         Gaps: 2
Percent Similarity: 65.625      Percent Identity: 64.583

alignment_block:
US-09-245-198a-4 x AC016876/rev ..

Align seg 1/1 to reverse of: AC016876 from: 1 to: 190358

159 AAGLYValaspolyThrValSerGlyTPGluAlaArgIleAsnse 175
:::|||||
83375 TCAGTGTGACGCGAGACGTGAGTGGCGGAGGAGCCAGATCAACAG 83326

175 rSerSerProLeuArGlyrAsnArgGlnIleGlyGluPheIleValThra 192
|||||
83325 CTCACACCCCTCTGCTGCTACCAACCGCCAGATCGGGAGATTATGATCACCC 83276

192 rGlaGlyLeuArGlyrLeuTyfCys..Gln..... 201
|||||
83275 GGGCTGGGCTCTACTACTCTACTGTCAGGTAGCCCATCTGCTGCAT 83226

201 ..... 201

83225 GGCTAACGACGTAAAGAGTGGCGAAGGTTTCCAGAGAGTGGGGGAC 83176

201 ..... 201

83175 AAGCTACAGGGCTGGAGAGGCTGAGTTGGGGTTTGGGTGGATGCC 83126

201 ..... 201

83125 TCGCTGGCTGAGAAATTGGAATTGAGCGGAGCGAGGAGGCTCG 83076

202 ..... ValHisPheAspGluGlyLysAlaVal 210
ACTGCGCTGTGTGCCACCCAGCGTGCACCTTGTGATGAGGGAAGGCTG 83026

210 alyrLeuLysLeuAspLeuValaspGlyValLeuAlaLeuArgCys 226
|||||
83025 TCACCTGGAAGCTGAGCTGCTGCTGATGCTGCTGCGCCCTGCGCTGC 82976

227 LeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuAr 243
|||||
82975 CTGGAGGAATTTCTACGCACTGCGCGCAGTCTTCTGCGGCCCAAGCTCCG 82926

243 gLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuAr 260
|||||
82925 CCTCTGCGACAGTGTCTGCTGCTGTTGGCCCTGCGGCGGAGGCTCCCTCC 82876

260 rglLeuArGlyrLeuProTAlaAlaHisLeuLysAlaAlaProPheLeuThr 276
|||||

```

```

82875 GGATCCGACCCCTCCCTGGGCCCATCTCAAGCTGCCCCCTTCTCACC 82826
277 TyfPheGlyLeuPheGlnValHis 284
|||||
82825 TACTTTCGACACTCTTCCAGGTTTCAC 82802

seq_name: gb_hhg:AC069459

seq_documentation_block:
LOCUS      AC069459      203083 bp      DNA      linear      HTG 27-JUN-2001
DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
SEQUENCE   SEQUENCE, 7 unordered pieces.
ACCESSION  AC069459
VERSION    AC069459.23 GI:14547768
KEYWORDS   HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 203083)
Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
Dederich,D., Thomas,S., Okunonu,G., Carllock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
Bunay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,J., Liu,M., Louisaged,H., Lozano,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mel,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Ogul,M., Parish,B.,
Perez,L., Reltter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Munhy,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,G.,
Worley,K. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 203083)
Worley,K.C.
Submitted (31-MAY-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Bay Jun 25, 2001 this sequence version replaced gi:12621364.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: MAFO
Center clone name: RP23-168P5
----- Summary Statistics
Sequencing vector: M13, L08821
Chemistry: Dye-Primer Bodipy: 48% of reads
Chemistry: Dye-Terminator Big Dye: 52% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210656; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 7.2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```



```

66609 TCAGGCTGATGGAGACAGTACGCTGCGAGACACCAAAATCAACAG 66560
175 rfsersrptoleuArigtYrAsnArgInlleglyglnpelleValpHra 192
66559 CTCACCCCTCTGCGCTACGACCGCAGATTGGGAAATTACAGTATCA 66510
192 rgaIaGlyLeuTyTyTLeuTyCys.Gln..... 201
66509 GGGCTGGGGCTACTACTGACTGTACAGTAAGCCCTGCGCTTCAG 66460
201 ..... 201
66459 GGGCAGAGCAAAAGCTAAGGGAGAGGAGACTGCGAAGAAATGGTGGA 66410
201 ..... 201
66409 GAGGAGACTTCGCGACTCATGAGAGAGATGCTTGGCTTCATGAGG 66360
202 .....ValHisPheas 205
66359 AGGCAGAGCTGTGATTTGCTTCCTCTCTCTTCAGGTGCACTTTGA 66310
205 pGIuGlyLySaIaValTyTLeuLysLeuAspLeuValAspGlyVal 222
66309 TGAGGGAAGGGCTGTACTACTGAGCTGACTGTGTGAGAGGCTGTCG 66260
222 euAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIAserSerLeu 238
66259 TGCCCTCGGCTGCTGAGAAATCTCAACCCACAGCAGCAACCTCTCCT 66210
239 gLyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgp 255
66209 GGGCCCCAGACTCGTTGTGCGAGGTGTCTGGGCTGTGGCGCTGGCGCC 66160
255 oGlySerSerLeuArgIleArgThrLeuProTTPAlaHisLeuLysAla 272
66159 AGGGCTTCCTTCGGATCGACCCCTCCCTGGCTCACTTAAAGCTG 66110
272 laProPheLeuThrTyTyrPheGlyLeuPheGlnValHis 284
66109 CCCCCTTCTCAACTACTTTGGACTCTTTCAAGTTCAAC 66072

seq_name: gb_hlcg:AC098923
seq_documentation_block:
LOCUS AC098923 179030 bp DNA linear HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS
ACCESSION AC098923
VERSION AC098923.4 GI:17973852
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 179030)
Munry,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimagne,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bovle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flaeg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 COMMENT

```

Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homi,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J.,
Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W.,
Loulisege,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabab,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wlasczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
2 (bases 1 to 179030)
Unpublished
Direct Submission
Worley,K.C.
Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064677.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G10K
Center clone name: CH230-154B15
----- Summary Statistics
Assembly program: Phrap; version 0.990329first call to
findPhrapList
Consensus quality: 137096 bases at least Q40
Consensus quality: 145080 bases at least Q30
Consensus quality: 152950 bases at least Q20
Estimated insert size: 138980; sum-of-contrigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contrigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 68 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 7467: contig of 7467 bp in length
* 7468 7567: gap of unknown length
* 7568 12115: contig of 4548 bp in length
* 12116 12215: gap of unknown length
* 12216 20313: contig of 8098 bp in length
* 20314 20413: gap of unknown length
* 20414 25589: contig of 5176 bp in length
* 25590 25689: gap of unknown length
* 25690 31434: contig of 5745 bp in length
* 31435 31535: gap of unknown length
* 31535 37595: contig of 6061 bp in length
* 37595 37596: gap of unknown length

```



```
* 37696 42172: contig of 4477 bp in length
* 42173 42272: gap of unknown length
* 42273 46334: contig of 3962 bp in length
* 46335 46334: gap of unknown length
* 50605 50605: contig of 4271 bp in length
* 50706 55202: contig of 4497 bp in length
* 55203 55302: gap of unknown length
* 55303 58407: contig of 3105 bp in length
* 58408 61286: gap of unknown length
* 61287 61386: gap of unknown length
* 61387 64567: contig of 3081 bp in length
* 64568 64568: gap of unknown length
* 64569 68498: contig of 3931 bp in length
* 68499 68599: gap of unknown length
* 73450 73449: contig of 4851 bp in length
* 73550 77402: contig of 3853 bp in length
* 77403 77502: gap of unknown length
* 77503 81027: contig of 3525 bp in length
* 81028 81127: gap of unknown length
* 81128 84144: contig of 3017 bp in length
* 84145 84244: gap of unknown length
* 84245 86680: contig of 2436 bp in length
* 86681 86780: gap of unknown length
* 86781 89210: contig of 2430 bp in length
* 89211 89310: gap of unknown length
* 89311 92339: contig of 3029 bp in length
* 92340 92439: gap of unknown length
* 92440 94951: contig of 2512 bp in length
* 94952 95051: gap of unknown length
* 95052 98586: contig of 3535 bp in length
* 98587 98686: gap of unknown length
* 98687 100597: contig of 1911 bp in length
* 100598 100697: gap of unknown length
* 100698 102766: contig of 2069 bp in length
* 102767 102866: gap of unknown length
* 102867 105237: contig of 2371 bp in length
* 105238 105337: gap of unknown length
* 105338 107305: contig of 1968 bp in length
* 107306 107405: gap of unknown length
* 107406 109057: contig of 1652 bp in length
* 109058 109157: gap of unknown length
* 109158 111454: contig of 2297 bp in length
* 111455 111554: gap of unknown length
* 111555 113704: contig of 2150 bp in length
* 113705 113804: gap of unknown length
* 113805 115933: contig of 2129 bp in length
* 115934 116033: gap of unknown length
* 116034 118193: contig of 2160 bp in length
* 118194 118293: gap of unknown length
* 118294 120414: contig of 2121 bp in length
* 120415 120514: gap of unknown length
* 120515 122311: contig of 1797 bp in length
* 122312 122411: gap of unknown length
* 122412 125628: contig of 3217 bp in length
* 125629 125728: gap of unknown length
* 125729 128545: contig of 2817 bp in length
* 128546 128645: gap of unknown length
* 128646 130563: contig of 1918 bp in length
* 130564 130663: gap of unknown length
* 130664 132945: contig of 2282 bp in length
* 132946 133045: gap of unknown length
* 133046 134556: contig of 1511 bp in length
* 134557 134656: gap of unknown length
* 134657 136478: contig of 1822 bp in length
* 136479 136578: gap of unknown length
* 136579 138159: contig of 1581 bp in length
* 138160 138259: gap of unknown length
* 138260 140456: contig of 2197 bp in length
* 140457 140556: gap of unknown length
* 140557 142544: contig of 1988 bp in length
```

```
* 142545 142644: gap of unknown length
* 142645 143784: contig of 1140 bp in length
* 143785 143884: gap of unknown length
* 143885 145304: contig of 1420 bp in length
* 145305 145404: gap of unknown length
* 145405 147391: contig of 1987 bp in length
* 147392 147492: gap of unknown length
* 147493 148908: contig of 1417 bp in length
* 148909 149008: gap of unknown length
* 149009 150147: contig of 1139 bp in length
* 150148 150247: gap of unknown length
* 150248 151752: contig of 1505 bp in length
* 151753 151852: gap of unknown length
* 151853 153441: contig of 1589 bp in length
* 153442 153541: gap of unknown length
* 153542 154953: contig of 1412 bp in length
* 154954 155053: gap of unknown length
* 155054 156244: contig of 1191 bp in length
* 156245 156345: gap of unknown length
* 156346 158095: contig of 1751 bp in length
* 158096 158195: gap of unknown length
* 158196 159309: contig of 1114 bp in length
* 159310 159409: gap of unknown length
* 159410 160682: contig of 1273 bp in length
* 160683 160782: gap of unknown length
* 160783 162602: contig of 1820 bp in length
* 162603 162702: gap of unknown length

alignment_scores:
    Quality: 385.50      Length: 176
    Ratio: 3.779      Gaps: 2
Percent Similarity: 57.955      Percent Identity: 51.136

alignment_block:
US-09-245-198a-4 x AC098923/rev ..
Align seg 1/1 to reverse of: AC098923 from: 1 to: 179030

160 GYVALAaPGlyThVAlSerGIYTrPGluGuaAlaArgIIeAsnSense 176
|||||
26419 GGTGTGATGGACAGTGTGAGGAGACCAAAATCAACAGCTC 26370
|||||
176 rseProLeuArgTYrAsnArgGlnIIeGIYluPheIIeValTrrArg 193
|||||
26369 CAGCCCTCGCCCTATGACCCGACATTTGGGAATTTACGGTCATCAGG 26320
|||||
193 IaGIYleuTYrTYrLeuTYrCys..... 200
|||||
26319 CTGGGCTCTACTACTCTACTGTACGTAAAGCCTGGGCTCCATGGTGA 26270
|||||
200 ..... 200
|||||
26269 AGCGATGCTAAGGGAGAGAGCTTGGCAAGAAATGGGTGGAGTGGGA 26220
|||||
200 ..... 200
|||||
26219 GAACCTGGGTTTCATGAGAAGATGCTGATTTCCATGAGCGACGCA 26170
|||||
201 .....GlnValHisPheAspGlu 206
|||||
26169 GAGGTCTGGAATTTGCTGTCTCTCTGTCACGCCAGTCACACTTATGAG 26120
|||||
207 GlyYsAlaValTYrLeuLysLeuAspLeuValaAspGIYValLeuAl 223
|||||
26119 GGGAGGACACTTACTTAAGCTGAGCTGTGGTGAATGGTGTCTGGC 26070
|||||
223 aLeuArgYsLeuGIuGIuPheSerAlaThrAlaIaSerSerLeuGIYp 240
|||||
26069 CCTGACACGCTCGGAAATATATACGCCACATGACGAGATCTCTGGGC 26020
|||||
240 rGlnLeuArgLeuCysGlnVal.SerGIYleuLeuAlaLeuArgProGI 256
|||||
26019 CCCAGATCCGTTTGGCCAGGCCGCTGTGGCTGTGTGATCCGCGGCCGG 25970
|||||
```

256 ySerSerLeuArgIleArgThrLeuProTrrAlaHisLeuYsAlaAlaP 273
|||||
25969 GCTTTCACCTTTCAGACCCGACCCCTTAACCCCTTAACCCCTCCG 25920
273 roPheLeuThrYrrPheGlyLeuPhe 281
|||
25919 CCGTCTATCCCGCCCGGACCTCTT 25894
seq_name: gb_pl:Ap002482
seq_documentation_block:
LOCUS Ap002482 187835 bp DNA linear PLN 10-JUN-2000
DEFINITION Oryza sativa genomic DNA, chromosome 1, clone:p0706B05.
ACCESSION Ap002482
VERSION Ap002482.1 GI:8468009
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (cultivar:Nipponbare) DNA, clone:p0706B05.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretidoideae; Oryzaceae; Oryza.
1 (bases 1 to 187835)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:p0706B05
Published Only in Database (2000) In press
2 (bases 1 to 187835)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (07-JUN-2000) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp,
URL:http://www.dna.affrc.go.jp:82/, Tel:81-298-38-7441,
Fax:81-298-38-7468)
The orientation of the sequence is from T7 to SP6 of the PAC clone.
Genes were predicted from the integrated results of the
following:GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as
SplicePredictor (October 1998 version). The genomic sequence was
searched against the non-redundant database NRP (PIR, SWISSPROT,
GENPEP, PDB) from MAF DNA bank and the cDNA sequence database at
RGP. Protein similarities of the coding regions were searched
against NRP with BLASTP2.0. ESTs represent the identified cDNA
sequences using BLASTN2.0 with the corresponding DBJ accession no.
and RGP clone ID.
Detailed information on overlap and assembly quality together with
annotation of this entry at
http://www.dna.affrc.go.jp:82/genomicdata/Genomefinished.html.
Location/Qualifiers
1. 187835
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="p0706B05"
1765..3195
/note="Similar to Streptomyces clavuligerus isopenicillin
epimerase (P18549)."
/codon_start=1
/protein_id="BAA96610.1"
/db_xref="GI:8468010"
/translation="MASLQSGDAANAVDADVDGAASPPSAPKRRACAAATYDAE
VRAPEAHDRVARLNNGTFCGCPASVIAARARQRLFLSDPAFYFHHLOPGLARSR
AAVAAVAGCAGSESLVDNYTAAITIMQVANSFARGDFARGLVLMFYTCSIK
NSIHAYARAGATYEVPLPVPSPDAIVAEFRALVARDGGRRRLAVIDHTA
MPVALIPRKELVAICREGVAKFVNDAAHAGVAVDRIADGETYASNLKWFECPS
AVAFIPIRKDDPVSKLHPVSSSEYNGSLMESAMIGVDYSQOLVVPDVAEFVNR
DGVEGIRRRHDKVEMGTMLAAAGTFLCTPPEMGCSMLVGLPGSLGVGSEDDAV
GLRTMLRKQFVEVPLVYNSKAAADAPPEVKNQDNPVTGYAIRISHQVYVNEEYE
ALRAVAKLVADGFCRLKRPKYNH"
join(6984..7174,7451..7576,7735..7846)
CDS
/note="hypothetical protein"
CDS
/codon_start=1
/protein_id="BAA96611.1"
/db_xref="GI:8468011"
/translation="MVVFPFAHPIHGTHHNNSTPGVATEGGGSGSSGKSMADAE
IDRRRRARRRRRPREPFIADGTLDPSQASRLYSIGFGSNFPIQLIEKYYTPICAFSONA
KVAREASRFSTIAVYMAIIDDCAQDSQKFLVTKHLFFWT"
complement(join(8422..8775,8875..9168,9253..9516,
9642..9908,9986..10830,10934..11207,11317..11842,
12632..12870,12993..13214,13312..13487,13585..13981))
/note="Similar to Arabidopsis thaliana chromosome 2, BAC
FLM4; putative ABC transporter (AC004411)"
/codon_start=1
/protein_id="BAA96612.1"
/db_xref="GI:8468012"
/translation="MPESMRDAEANASSASVANAADSPGNGKGGGGAAGVARGER
AASASARVPEFHLKFAADKTDALVAALFTGLVANGALPMPFVLFGLIAPFGA
MGHIDVYNRMSMSEFTYLAIASVAFPQVYTCMTTGERQARINRLKLTILRE
IAFDKTYNTGEVGRMSGDFVLIDQAMGEVKFIDLVTFGLGFTVAQOGLTL
VMAATIPPLVAVGAVMSNVYAKMASLQAAVAESVVEQDTIGIRIVASFTEGKQAV
EKYKSLKSAVSGSREGLAAGLGMGVVWLLFCGYSLGIWYGAKLILLYGAKVM
NIVRAVLGSLALGQASPPSKAFAGGQAAVYKMETINRRPEIDAYSTGTMKPDIDG
DIERFDVYFSPTRPDDOIFRGFSLTDSGTVALVNGSGSKSTVSLIERFDPL
GDVILDEVNKEPOLRINRKTICGVSEPVLPAASTIKENTATKGDNATDDEIRAAAL
ANASKFDKMPQGLDTSVGEHTQLDSGQORAIARAILDPRILDLDETSALDE
SERIVQALDRVMNRTTYIVARLSTVRNADTAVIHQGLVERGPHBLDPEGA
YSOLIKLOEANRQDSDRKGDSGARSKQSLINSQASRSRSDSHSFSVPEGM
LGIDIOGSDNLCDGMPDVPVLSRLASLUNKPEIPVLILGSIASVIGVFPIAIL
SNVIAKFEPEPHILRKDSQFMSMEIYFVLSLTPVSSYLSIAGCRILIKRILAT
FEKYNNETIEFHPENSSGATGRLSADAKVAGLYGDALQALQVQNTLLAGLYA
FVSNWELIILALPLILGLNGWITOMFTIQGFSADAKMTEASQVANDAVSIRTY
SFSAEKVMYDLYKRCGPLETGTIRGTIISQIGVSEFLLFVYASFAAGALVE
NKTTFPVPYFRVFLALMAAIGVQSSTLITSSKAKSAVSIETIVRSKSIDSEDA
GVTYETLHNGIIEFQHVSEFRYPTRDVEIFRDLCTIHSGKALVAGSGSKSTAIL
LQRFYDDVGHILIDGVQIDKFLKMLROOMGLTISPKALFENDVRANIVYKGEAT
ESETIEPAKLANAKRTISSHQGCTTVCERGADLSCGQORITAIAYIDPKILL
DEATSAIDAESERVQDLDRVMNRTTYIVARLSTIQNDLAIYKNGVITIEKGNH
DTLNNIKDGAVALVAHLSAASS"
join(16079..16164,17463..17646)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96613.1"
/db_xref="GI:8468013"
/translation="MAHTHRGACVRSVARGVSAQDSCAGCDVDAIGPCVT
SELFYSYRKHSKIKPCPNRGRDLPAGNIGQGRREIYHEHNSN"
join(19104..19259,19684..19745,19879..20015,20087..20140,
21463..21598,24585..24645)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96614.1"
/db_xref="GI:8468014"
/translation="MOYCAVAVRYARVANVSSSRKQPCGRSIDVTLMTTNNLSP
LRPDTGCVDPHGATASRGDRDRDRTYARAHEHGCGRKREISRRRRRVRISLR
TOVYSHMIDLRTHNEGDYIYHHEHYIISRGVEVPEPTALTDEFTCPSSFLIYVR
ILSGDLIVERCHINSHSAPLVIESIGSNVYANRQFT"
complement(join(27285..27309,29006..29087,29375..29474,
29539..29601))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96615.1"
/db_xref="GI:8468015"
/translation="MTRAAVYHLNTSSSTAGCLAAMLGRGICQWRMSRDEQAASSIE
AMDLASGAGMALKRKGKKEIKMLCSKSCAPEVASRRMPTTRKM"
complement(30314..31296)
/note="3' UTR"
join(31975..32199,32416..32620,32656..33163,33648..33771)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96616.1"
/db_xref="GI:8468016"
/translation="MPEDTVARAGESRRPRYPERSYHLLSFAMAGLIPPSREFHY
IDFETIHALHAPNAVHTLAFVILCEFIETRSYAMLAVSDLGCEYAVDRRLGLN
QGLTGAVFGDYFHCRIAPLOERSRGMEYTGPDPMKTHVDGSLAAGSGHRRAND
SRMDPILCSDRDRESILAVMTVAGRGRRSGAGGGGDTGTSSGATAGGGRNGSS

```
CDS
      GGDGSRAPGPNRPGGDSASDPKRRKMSFPPSPPHGGAERTADRPAGHKCPAA
      SGVGRKKRLRKIGTEPCRCGSFIEPPKTFKRPDRDPLSGSTDESGSEGG
      SDDMDVVGAGIGLGP"
      complement(join(34469..36351,36628..37001))
      /note="Similar to Oryza sativa GAG-FOL precursor
      (AB030283)"
      internal stop codon: complement(36927..36929
      probably inactive because stop codons and frameshift
      positions are included in CDS"
      /codon_start=1
      /pseudo
      /db_xref="GI:34613
      /note="one bp frameshift insertion"
      /db_xref="GI:34925
      /note="one bp frameshift deletion"
      /db_xref="GI:35794
      /note="one bp frameshift deletion"
      /db_xref="GI:40163
      /note="5' LTR"
      complement(join(41162..41301,41962..42181,44317..44345,
      44535..44778,46639..46732,47583..47695))
      /note="hypothetical protein"
      /codon_start=1
      /protein_id="BAA96617.1"
      /db_xref="GI:8468017"
      /translation="MLTTTIDRVNRVHCRFIDLGEFPLATMHOQOVSIRKIRVDY
      VERRPATICMDPFEFVGMILDMRERAAAGSVSGTATANERPSPCALPPM
      TMRRRGVAPSPSTGATNDDEEEEARASGATSDKNAASVAALIVTCRCILM
      GSKMASVITLGAAGSNIVFMVLSDSVVLAVHRIISDGTGASRESEFAOASIALRL
      HHHAAPCCLRPNQARETPADRRKSEFFILSTYGQCRPIEDGKKMGHDLVCLV
      FYN"
      join(48744..48975,49715..49749,49993..50176,51131..51258,
      51368..51468,51518..51630,51771..51915,52454..52930,
      53116..53324,53464..53712,53894..53976)
      /note="hypothetical protein"
      /codon_start=1
      /protein_id="BAA96618.1"
      /db_xref="GI:8468018"
      /translation="MLAGSDHHRAGPRRGGRHCRRRVGEKMRMPRDAREKROPHM
      DEVPGLMPEMIVLEFVLAESSGRHARLRGRLGEDNDCKSLSVGOTMRRKVN
      GDITKNTMSSTLPPPIRPDKIVFSSLMFGGFPDEMLGEBGLVSAMYNLALSTWA
      RAKAVLILGYSTIESIRSLVSPFTEVPSONKVLQRIIDVANEKRNKKTWIDHW
      SYTKNKKRMAAGAGGAGVQVPGDKGLAAROGGSGMSACPCACANPLPPAA
      HRFESLALPVPADPNALPLSLPVGRRALPLISRSVGRGVALETERRRRRAGEG
      RSGGGDRDRRLAVGRKTHONGRLVATGPAIALSTVLHALSAGHDAASAVL
      PRPRCKGRRLATATRYGIRSKMAASGGGFCGCGGGVGGRSGAGCGGAADG
      KRRRAPANDVYTRYLGICMHTAMSSSPYVSIVSYFVGNIRKEKEQVRYPEAA
      GADGVADRLLAAAPELSTFAKTKETTVAAAGAAARSLKQPRHPRHLGFEES
      AADDISMPAVGCGMWMCCRRNGERMDSSLPLPPLKLPALAIASALARRSPLADG
      ROYRHISATSAKPSKTAEGVKLHRFO"
      complement(54260..54721)
      /note="ESTs AU031727(R0930), D24036(R0930) correspond to a
      region of the predicted gene.
      hypothetical protein"
      /codon_start=1
      /protein_id="BAA96619.1"
      /db_xref="GI:8468019"
      /translation="MAGAAMPAPAPAGNGGFCNCPYPSVVAESGAVEAADVAG
      VYGMIREAKARLPADPARAVAVDFREDPELSTDSLSVSMTRIGLEADGKQPP
      SHVIPPAYTFPAVAATIGAPVPVKGTGARIMTCVEEDHLPAREKEELKADK"
      complement(57466..57918)
      /note="EST AU032068(R3431) corresponds to a region of the
      predicted gene.
      Similar to Nicotiana tabacum mRNA for HSR201 protein.
      hypersensitivity-related gene (X95343)"
      /codon_start=1

alignment_scores:
  quality: 145.50      length: 314
  ratio: 0.983      gaps: 15
  percent similarity: 47.134      percent identity: 26.752

alignment_block:
```

```
us-09-245-198a-4 x Ap002482 ..
Align seg 1/1 to: Ap002482 from: 1 to: 187835

13 LeuProLeuPro...ArgSer.LeuGlySerArgAspGlyAlaVala 28
|||||
32532 CTTCCATTGGCCGACATCGCACCCCTCCAGAGACGGCTCCGTCGCCATGCG 32581
28 rglInlaGlnProProAlaProMetAlaAlaArgSerGlnArg 44
:::
32582 AGTACACTGGGCCACAGATCCATG.....CGCACCATGTGGGTGAG 32625
45 ArgGlyArgArgGlyGluProGlyThrAlaLeuValaProLeuAlaLe 61
|||||
32626 CGCTGGGACCTAGGGCCGAGAGAGCGCGCAAGATGATGATCTGCCAGTCT 32675
61 uGlyLeuGlyLeuAlaLeuAlaCysLeu.....
32676 GGGTCTGGACATCCCGCAGCAACACATGATTCAGATGAGATCTCCCT 32725
70 .....
32726 CTGACGCGACCGGACCGGAGAGATCTCCGCTGATGACGTCCTCG 32775
71 .....GlyLeuLeuValaValaSerle 79
32776 GCGCTGGCAGGGGTCGCTCCCGCGGACGCGCTGCGGTGTCGCGAT 32825
79 u.....GlySerArgAla 84
32826 GGCACGGGACACAGTGTGCGACCCGAGTGGCGGATGCGATGCGGCTC 32875
84 erLeuSerAlaGlnGluProAlaGlnGluGluLeuValaAlaGluGluasp 100
32876 CAGCGGGGCGACGGCTCCAGGCGCCGCGTCTGATTCGTG..... 32916
101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAl 117
32917 ..GCCCGGGGAGATTCGCCGACGACCGCAAGAAAGCAAGCAAAATG 32963
117 apProHeuAsnArgLeuValaArgProArgSerAlaProLysGly 134
:|||||
32964 TCCAGATTGGCGCGCTTCACCCAGCATGAGAGGGGCGCGAGCGAAC 33013
134 rglVSThrArgAlaArgArgAlaLeuAlaAlaHISYrGluValaHISPro 150
GCGGACCGACCCACCGGGGCGCCAAATGCC.....CTGCGC 33051
151 ArgProGlyGlnAspGlyAlaGlnAlaGlyValaasp..... 162
33052 CGTCCGGGGTGGAGCGGAGAAAGCGGCTCCGGAAGATAGGCAAAAGC 33101
163 .....GlyThrValSerGlyTTPGluGluAlaArgIleasnSers 176
33102 GAGCGGTGCGGGGAGTTCATCGAGCCGCCCAAGTGA...CCTCAA 33148
176 erSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArg 192
33149 AGCGCCCTC.....GCAGGTGAGATTGCCAGATCCAA 33183
193 AlaGly...LeuTyrTyrLeuTyrCysGlnValaHISpAspGluGly 208
|||
33184 CCGGTCGCTTCATCCACTCGGTTTCCTTTTTCCTTCCTTCCTTT 33233
208 sAlaValTyrLeuLysLeu.....AspLeuLeu.....Vala 219
:|||||
33234 CTCTCTTTTAACTTGATCGGGGCTCTCTAGTAGATTCATCCATCC 33283
219 spGlyValLeuAlaLeuArgCysGlnGluGluPheSerAlaThrAlaAla 235
|||
33284 CAGGGGCTTCGACGTCGATCCCTCACAGGGGTGAATTCGAGATC 33333
236 SerSerLeuGlyProGlnLeuArgLeuGlyValSerGlyLeuLeuAl 252
```

```
33334 GGCATCGCTTGTCCTCCGCTGATCAGATCCAT.....GGAGTCTCTCCG 33377
252 aleuArg.....ProGlySerSerLeuArgIleArg 262
33378 AGCATCTCTGGCGGGGCCACGGCCGCGGTCGG 33417
```

seq_name: gb_ba:SC1C2

seq_documentation_block:

LOCUS SC1C2 42210 bp DNA linear BCT 15-JAN-1999
DEFINITION Streptomyces coelicolor cosmid 1C2.
ACCESSION AL031124
VERSION AL031124.1 GI:3355667
KEYWORDS 3-isopropylmalate dehydratase large subunit; 3-isopropylmalate dehydratase small subunit; 3-isopropylmalate dehydrogenase; branched-chain amino acid aminotransferase; carboxyl transferase; delta-1-pyrroline-5-carboxylate dehydrogenase; glx; glutamyl-tRNA synthetase; histone-like DNA binding protein; hydrolase; lve; leuB; leuc; leuS; lyase; secreted lyase; transfer-RNA-Gln; transfer-RNA-Glu; ureaB; urease alpha subunit; urease beta and gamma subunits; ureC.

SOURCE
ORGANISM Streptomyces coelicolor A3(2).

REFERENCE
AUTHORS Streptomyces coelicolor A3(2).
JOURNALS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
REFERENCE Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
AUTHORS 1 (bases 1 to 42210)
JOURNALS Murphy, L. and Harris, D.
REFERENCE 2 (bases 1 to 42210)
JOURNALS Parkhill, J., Barrell, B. G. and Randal, M. A.

REFERENCE
AUTHORS Direct Submission
JOURNALS Submitted (22-JUL-1998) Streptomyces coelicolor sequencing project,
REFERENCE Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
AUTHORS CB10 ISA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
JOURNALS David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
REFERENCE Colney, Norwich, Norfolk NR4 7UH, UK
JOURNALS 3 (bases 1 to 42210)
REFERENCE Redenbach, M., Kleiser, H. M., Denaplatte, D., Eichner, A., Cullum, J.,
JOURNALS Kinashi, H. and Hopwood, D. A.

REFERENCE
AUTHORS A set of ordered cosmids and a detailed genetic and physical map
JOURNALS for the 8 Mb Streptomyces coelicolor A3(2) chromosome
REFERENCE Mol. Microbiol. 21 (1), 77-96 (1996)
JOURNALS 97000351

COMMENT

Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.

Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

(URL: <http://www.sanger.ac.uk/projects/S-coelicolor/>) CDS are numbered using the following system eg SC787.01c. SC (S. coelicolor), 787 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an upstream codon (atg, gty, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence

FEATURES
source overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 1C2 lies between 8D9 and 7A1 on the AseI-B genomic restriction fragment.
Location/Qualifiers

source

1..42210

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100222"

/clone="cosmid 1C2"

1..1228

/gene="SC1C2.01"

<1..1228

/gene="SC1C2.01"

/note="SC1C2.01, probable delta-1-pyrroline-5-carboxylate dehydrogenase, partial CDS. len >408 aa; similar to many e.g. TR:050443 (EMBL:AL010186)

delta-1-pyrroline-5-carboxylate dehydrogenase (M. tuberculosis) (543 aa), fasta scores: opt: 1753 z-score: 2214.2 E(): 0, 65.1% identity in 407 aa overlap, and pun2_HUMAN delta-1-pyrroline-5-carboxylate dehydrogenase (563 aa), fasta scores: opt: 1273 z-score: 1730.4 E(): 0, 48.5% identity in 408 aa overlap. Contains PS00687 and PS00070 Aldehyde dehydrogenases glutamic acid active site and cysteine active site and Pfam match to entry PF00171 aldedh. Aldehyde dehydrogenases, score 114.70, E-value 1.8e-30"

1..3807

/codon_start=2

/transl_table=1

/product="delta-1-pyrroline-5-carboxylate dehydrogenase"

/protein_id="CAI1968.1"

/db_xref="GI:3355668"

/db_xref="SPRMBL:086502"

/translation="IDSPCELDIFMRNVHVARNIIAEPANSGVNMRRDRPLEG FVAATPFNSALIAANLPTAPALMGVNVKPSPTQHAVALMLQLEAGLPGVNL LMGDAIVSEVLAHEHDLGIHTGTSKTFKQHLMKVGNIRKRYRPLAETGK DLVAPHASARAVIKTRALTGARGFYQGSATSRATIPASINDDGKEFAAEVTL TMGDVTLDSNFTGAVIDERSFAKRAIDRAKDECTIVAGSDYDSVFEVPTV ECDPENEEVRETEYFEPFLAVHYVDSADADAYAMLIQMSVSDYALTGVSINDRA AATYMEKLRVAAAGNFYINDKSTGAVVGQPFGGGRASGTNDKAGAPQNLKRWTLTRAI KETLVAPDTYTYPMG"

1..3807

/note="true overlap with cosmid 8D9"

2..1189

/gene="SC1C2.01"

/note="Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenases, score 114.70, E-value 1.8e-30"

464..487

/gene="SC1C2.01"

/note="PS00687 Aldehyde dehydrogenases glutamic acid active site"

548..583

/gene="SC1C2.01"

/note="PS00070 Aldehyde dehydrogenases cysteine active site"

1246..1283

/note="hairpin_loop with 18bp stem"

1522..1527

/note="possible RBS upstream of SC1C2.02"

1537..1668

/gene="SC1C2.02"

1537..1668

/gene="SC1C2.02"

1537..1668

/note="SC1C2.02, questionable ORF, len: 43 aa"

/codon_start=1

/transl_table=1

/product="hypothetical protein SC1C2.02"

/protein_id="CAI1969.1"

/db_xref="GI:3355669"

/db_xref="SPRMBL:086503"

/translation="MAETTVRRRVRLHLSRTSDSRKNAALQSLDRDNGATGH"

1744..2787

/gene="leuB"

1744..2787

CDS


```

* 226315 227488: contig of 1174 bp in length
* 227489 227588: gap of unknown length
* 227589 228852: contig of 1264 bp in length
* 228853 228952: gap of unknown length
* 228953 229963: contig of 1011 bp in length
* 229964 230063: gap of unknown length
* 230064 231343: contig of 1280 bp in length
* 231344 231443: gap of unknown length
* 231444 232645: contig of 1202 bp in length
* 232646 232745: gap of unknown length
* 232746 233912: contig of 1167 bp in length.
Location/Qualifiers
source 1..233912
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="Rfl"
/clone="CH250-192N7"

```

```

BASE COUNT 58900 a 57745 c 56429 g 58611 t 2227 others
ORIGIN

```

```

Alignment_scores:
Quality: 136.50 Length: 314
Ratio: 0.916 Gaps: 24
Percent Similarity: 47.452 Percent Identity: 31.210

```

alignment block:

US-09-245-198a-4 x AC105470/rev ..

Align seg 1/1, to reverse of: AC105470 from: 1 to: 233912

```

3 LeuLeuAspPheGluLeuIleSerAlaArgArgLeuProLeuProArgSerLeu 19
|||||
32662 CTCTCTCAT.....AGCGTAGGGGCTGCG...ACTAGAGCGAT 32625

19 uGlySerArgASP.....GlyAlaValAlaArgLna 30
:|||||
32624 AGGCTGTGTGAGACCGGACAGCCCTAGGGGGTCCGAGTGTGACCGCG 32575

30 lagInProProAlaProMetaAlaAlaArgArgSerGlnArgArgGly 46
|||||
32574 CCAAGCCCGGCTCGCTCTGCGGAGTAGAGCGCGCGCTCATCGG 32525

47 ArgArgGlyGlu...ProGlyThrAlaLeuLeuValProLeuAlaLeuG1 62
|||||
32524 CGCGGAGGTGAGCGCGCGGC.....CTCGG 32499

62 yLeu.GlyLeuAlaLeuAla...CysLeuGlyLeuLeuAlaValAl 77
|||||
32498 GCGGTGGCGGAGCTCGGGGAGTCCCGCGGGCTGACGACATGTGCTC 32449

78 SerLeuGlySerArg,AlaSerLeuSerAlaGlnGluProAlaGlnGlu 94
:::
32448 GCC.....GCTCGTGTGAGAGCGCGCTGCGCGTGGCGGACCATGTGCG 32405

94 lueuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThr 110
:::
32404 AC.....GCGGCTGAGAGAGCAG...CCCATGAGAGACGAGCGGCCACC 32364

111 GluGluSerGlnAsp.ProAlaProPheLeuAsnArgLeuValArgPro 127.
|||||
32363 GAGAAAGGACACGAGCGCGCGCCG.....AAGCGAGGCGCC 32326

127 rGArgSerAlaProGlyGly...ArgLysThrArgAlaArgAlaIle 142
|||
32325 GGTGAGGCTTACCTGGGCGAGCGCGCCCGCGCATCTCGCGGGGTCTG 32276

143 AlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAl 159
:::
32275 GGGGCG.....GGACCAAGCGCGCGCGCGGAGGACCAAGAAC 32238

159 aGlyValAspGlyThrValSerGlyTyrGluGluAlaArgIleAsnSer 176
:::

```

```

32237 GGCGCCGAGGCGAC.....AGAT 32218

176 erSerProLeuArgTyrAsnArgLniIleGlyLuphelleValThrArg 192
|||||
32217 CAAGCCAGCAAGAACGAGGAGACGCGGGT.....AGG 32183

193 AlaGly.....LeuTyrTyrLeuTyrCy 200
:::
32182 TGGCGCCCAAGCGCGCGCGCGCTTGTTCGCGGGAGTCTTTTG 32133

200 sGlnValHisPheAspGluGlyLys..... 208
|||
32132 TTGGCGCCACGCGCGGGGGAGGGGAGCGGCCCGCGAAGATGCT 32083

209 .....Ala 209
GGCGCGCTCGCGCGCGCGCGCGCGCGCGCTTGTGAGTTTGCA 32033

210 ValTyrLeuLysLeuAspLeuLeuValAsp..GlyValLeuAlaLeuArg 225
:::
32032 GGAGTTTGTGGCGTGCAGCTTGAAGCGCGACCGGGCTGCTCGCG...CGG 31986

226 CysLeuGluGluPheSerAlaThrAla..... 234
|||||
31985 TGCTTA...GGCTACAGCGGCTGCGCGCGCGCGCGCGCGCGAAGAGGAA 31939

235 .....AlaSerSerLeuGlyProGln..LeuArgLeuCysGlnVal 248
:::
31938 AAGCGCGGATGCGGATCGGATCGGCGCGCGCGCGCTGCTGTGCTGCGTAG 31889

248 erGlyLeuLeuAlaLeuArgProGlySerSerLeu 259
:::
31888 GCGGAGTTGGCGCGCGCGCGCGCGCGCGCGCTAGCTTG 31854

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 10:48:59 ; Search time 2447.42 Seconds
(without alignments)
7571.777 Million cell updates/sec

Title: US-09-245-198a-3
Perfect score: 1373
Sequence: 1 atcgtatgttgaacttga.....gacaaatgtataatg 1373

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
otal number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	765.8	55.8	834	10	BI766766 603056866
2	743.4	54.1	777	10	BI819200 603038614
3	701.4	51.1	731	10	BI871711 603395825
4	688	50.1	828	10	BI596681 603243254
5	677.4	49.3	728	10	BI870393 603395641
6	584.4	42.6	609	10	BI966060 1672904.x
7	581.4	42.3	1027	12	AF163779 AF163779
8	531.6	38.7	910	10	BG110063 602279667
9	516.8	37.6	531	10	BI824443 603038693
10	493.6	36.0	918	10	BI824443 603038693
11	458.2	33.4	1033	11	AK020909 Mms muscu
12	454.4	33.1	561	9	AW763237 ur70d09.y
13	442.8	32.3	587	10	BG686319 602638232
14	440.2	32.3	456	10	BI966255 1672904.y
15	436.8	31.8	440	10	BM128059 1109B06.y
16	431.6	31.4	538	10	BF821434 MRI-RT003
17	404	29.4	436	12	AQ890280 HS_3188_B

18	399.4	29.1	785	10	BI762908
19	394	28.7	413	9	AI422796
20	389.2	28.3	698	10	BI906850
21	374.2	27.3	416	9	AI291866
22	372.6	27.1	584	9	AW917574
23	365.4	26.6	422	10	BI677255
24	365	26.6	399	9	AI913541
25	363.4	26.5	367	12	AO100365
26	359	26.1	474	10	BI965174
27	358.4	26.1	407	9	AI221985
28	357.6	26.0	894	10	BI908274
29	352.8	25.7	409	9	AW131279
30	347.2	25.3	397	9	AI669243
31	342.6	25.0	710	10	BE858778
32	339.8	24.7	372	10	BI677256
33	327.8	23.9	650	10	BG404836
34	327	23.8	346	10	BE858822
35	327	23.8	471	9	AA221610
36	326	23.7	337	9	AW195034
37	326	23.7	345	10	BF439993
38	326	23.7	352	9	AM204512
39	326	23.7	364	9	AM291620
40	318	23.2	329	9	AI695776
41	318	23.2	340	10	BF195436
42	318	23.2	542	10	BF041509
43	317	23.1	346	9	AI760777
44	315	22.9	346	10	BF940141
45	314	22.9	351	10	BG054914

ALIGNMENTS

RESULT 1
LOCUS BI766766 834 bp mRNA linear EST 25-SEP-2001
DEFINITION 603056866F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
mRNA sequence.
ACCESSION BI766766
VERSION BI766766.1 GI:15758344
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1517 row: C column: 18
High quality sequence stop: 772.

FEATURES

location/Qualifiers
1..834
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_id="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size

Best Local Similarity	99.3%	Pred. No. 1,2e-124	Mismatches 1	Indels 4	Gaps 2
Matches 726	Conservative	0	Mismatches	1	Indels
OY	192	cctgagcgtgagcctgagcctcctctgctgacgtggtgagtcagtttgaggagccggacac	251		
Db	1	CCGTGGCGGTGGGCTGCTGCTGCGGCTCTGCTGCGCCGCGTCAATTGGGGAGCGGGCATC	60		
OY	252	gctgtccgcg---caggagagcctccagagagagctgtgtagagaggagagcagagacc	308		
Db	61	GCTGTCCGCGCCACGCAAGAGGCTGCCAGAGAGAGCTGGTGGCAGAGAGACAGGACCC	120		
OY	309	gtcgaagactgaaatcccccagacagaaagacaggatccctgagccttctgtaaccgact	368		
Db	121	GTCCGAACTGAATTCGCCAGAGAAAGAAAGCAGAGATCCTGGCGCTTCTGAACGACT	180		
OY	369	agttcgccctcgagaaagtgcacctaaaggccgagaaacaggggtcgaagagcgatgc	428		
Db	181	AGTTGGGCTCCCAAGAGTGCACCTAAAGGCGCGGAAACACGGGCTCGAAGGCGATCCG	240		
OY	429	agcccatatagagctatccagagcccgctggaagagacagagcgagagaggtgtgaagc	488		
Db	241	AGCCCATATATGAAGTTCACTCCAGACTGACAGACTGACAGAGGACGCGAGGAGTGTGACCG	300		
OY	489	gaacagtgaatgctgtaggaggaagccagaatcaacagctcacaagcctctgcgtacaaacc	548		
Db	301	GACAGTGAATGGCTGGGAGAGCCAGAAATCAACAGCTCACCCCTCTCGCTACAAACCG	360		
OY	549	ccagatcgagagattatagtcacccgggctggtctactactcgtactgtcaggtgca	608		
Db	361	CCAGATCGGGAGATTATATGTCACCCGGGCTGGGCTCTATCACTGTGTACTGTCAAGTGA	420		
OY	609	ctttaaagaggaagagcgtgtcctaacctgaagcctgtagactgtgcgtgtagatgtgtgctgac	668		
Db	421	CTTTGATGAGGGGAGAGGCTGTCTACTGTGAAGCTGGACCTTGCGTGGTAGATGATGCTGTGCG	480		
OY	669	cctgagcgtgctgtagagaaatctcaagccaatgctgagccagttccctcgaggcccaagctcg	728		
Db	481	CCTTGGGCTGCCGAGAGAAATTCACGCCACTCGGGCCACTTCCCTCGGGCCCCAGCTCCG	540		
OY	729	cctcgagcaggtgtctgtgagctgttgagcctctggcgcaaggtctccctcgatccgac	788		
Db	541	CCTCTGCGCAAGTGTGTGGGCTTTTGGCCCTGCGGCCA-6GTCTCTCCTTCGGATCCGAC	599		
OY	789	cctcccttgaggccatctcaagcgtgcccctctctcaactaattcgaactcttcaggt	848		
Db	600	CCCTCCCTGGGGCCATCTCAAGGCTGCCCCCTTCCTCACTACTTGGGACTCTTCCAGGT	659		
OY	849	tcacgaagaggagcctggtctcccccagatgctgtcccaaggtctcccgctccctccgaagct	908		
Db	660	TCACGAGAGGGGCTCGGTCTCCCGCAATGTCCTCCAGAGGCTGCGGGCTCCCTCGACAGCT	719		
OY	909	ctctgagcacc 919			
Db	720	CTCTGGGCGCC 730			
RESULT 4					
BI596681		828 bp	mRNA	linear	EST 07-SEP-2001
LOCUS					
DEFINITION		6032242254P1 NIH_MGC_96	Homo sapiens	CDNA clone IMAGE:5285892 5',	
ACCESSION		BI596681			
VERSION		BI596681.1		GI:15489620	
KEYWORDS		EST.			
SOURCE		human.			
ORGANISM		Homo sapiens			
REFERENCE		EMBL:U014, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE		1 (bases 1 to 828)			
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.			
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)			
		unpublished (1999)			
		Contact: Robert Strausberg, Ph.D.			

Email: cgarbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1722 row: k column: 13
 High quality sequence stop: 776.
 Location/Qualifiers

FEATURES

source
 1..828
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5285892"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptPR (modified
 pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 155 a 278 c 223 g 172 t
 ORIGIN

Query Match 50.1%; Score 688; DB 10; Length 828;

Best Local Similarity 95.3%; Pred. No. 4.6e-120; Matches 732; Conservative 0; Mismatches 30; Indels 6; Gaps 2;

QY 312 ggaactgaatcccccagacagaagaagccagatcctcgccttccctgaacccgactagt 371
 DB 43 GGAATGGAATCCCGACAG 102
 QY 372 tgggctcgcgaagatgacacttaagccggaagaaacagggctcgaagagcgatcgaagc 431
 DB 103 TCGGCTCGGAGAGATGCACCTAAAGCCCGGAAACACGGGCTCGAAGACGATCGAGC 162
 QY 432 ccaatgaagatccatccagacagcagagcagcagcagcagcagcagcagcagcagcagc 491
 DB 163 CCAATTATGAAGTTCATCCAGACACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
 QY 492 agtgaatgagctggaggaagccagaatcaacagctccagcctctgcgtacacaacgcga 551
 DB 223 ATTGAGTGGCTGGAGAGAACCCAGAAATCAACACACTCCAGCCCTCTGGCTATCAACGCCCA 282
 QY 552 gatcgggagatctatagtcacccgggctggcctactactactactactactactactact 611
 DB 283 GATGCGGAGATTTATAGTCACCCGGGCTGGGCTCTACTACTCTACTCTACTCTACTCTACT 342
 QY 612 tgaagaagggaagcgtctactgaagcgtgaagcgtgctgctgctgctgctgctgctgct 671
 DB 343 TGATGAGGGGAGAGGCTGTACTGAAAGCTGGAAGCTGCTGTGTGTGTGTGTGTGTGTGTGT 402
 QY 672 ggcctgcctggaggaattccagcactgcgcagcttccctcgcggccacagctccgct 731
 DB 403 GCGCTGCTGGAGGAATTCACAGCCACTGCGGCGAGTTCCCTCGGGGCCACAGCTCCGCCCT 462
 QY 732 ctgcagaggtctggtgctgtgtgctcctgcggcagaggtctctcctcctgaggtccgaacct 791
 DB 463 CTGCGCAGGCTGTGGCTGTGGCTCGGCGGCGGCGGCTCTCCCTCGGAGTCCGACACCTT 522
 QY 792 cccctggccacatccaaagctgccccctcccaactactcctcgaagcttccagagttca 851
 DB 523 CCGCTGGGCGCCATCTCAAGAGCTGCCCCCTTCTCCTACTACTTGGAGACTCTTCCAGGTTCA 582
 QY 852 ctgaaggggccctggtctccccaagtcgtcccaagctgcgcgctccctcgcagacgtctc 911

DB 583 CTGAGGAGGCGCTGTGCTCCCGCCAGTCGCCAGGCTGCGGCTCCCTCCGACAGCTCTC 642
 QY 912 tgggaccccggttccctctgcccacccctcagcagcgtcttctgctcagagccctgccccttc 971
 DB 643 TGGGACCCGCGTCCCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 702
 QY 972 ctctagaagctgctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1025
 DB 703 CTCTAGAGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 762
 QY 1026 attccacactctatcttaacactcccccacgcgcacactcctcaactca 1073
 DB 763 TTTCCCACTCTTATCTTACAACTTCCCGCCACAGACCAATCATTCACA 810
 RESULT 5
 BI870393
 LOCUS BI870393 728 bp mRNA linear EST 11-OCT-2001
 DEFINITION 603395641F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405459 5',
 mRNA sequence.
 ACCESSION BI870393
 VERSION BI870393.1 GI:16044066
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 728)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgarbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12034 row: a column: 12
 High quality sequence stop: 728.
 Location/Qualifiers

FEATURES

source
 1..728
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5405459"
 /clone_lib="NIH_MGC_90"
 /tissue_type="adrenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 125 a 240 c 227 g 136 t
 ORIGIN

Query Match 49.3%; Score 677.4; DB 10; Length 728;
 Best Local Similarity 99.0%; Pred. No. 4.9e-120; Matches 724; Conservative 0; Mismatches 1; Indels 6; Gaps 4;

QY 192 cctggagcgtgagcctgctgctcctctgctgctgctgctgctgctgctgctgctgctgctgct 251
 DB 1 CCTGGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 QY 252 gctgtccgccc---caggagcctgccagagagagcgtggtgagcagagagagagagagagagcc 308
 DB 61 GCTGTCCGCCACAG 120
 QY 309 gtccgaactgaatcccccagacagaagaagccagagatcctcgccttctcctgaacgact 368

```

|||||
Db 121 GTCGGAATGTAATCCCGAGACAGAAAGCCAGGATCTCGCTTTCGTAAGCCGACT 180
Qy 369 agtcgcctctgcgaagatgacactaaagccggaaaaacagggctcgaagaagcatgac 428
Db 181 AGTTGGGCTTCGAGAAATGACCTAAAGGCGGAAACACGGGGCTCAAGAGCCATCGC 240
Qy 429 agcccatatgaattcatcacaagcctgcgaagcagcagcagcagcagcagcagcagc 488
Db 241 AGCCATTATTAAGTTCTTCCACGACCTCGACAGGACGACGACGACGAGGTGAGAGG 300
Qy 489 gacagtgagtgctgcggaggaagcagaatcaacagctccagccctgcgctacaacg 548
Db 301 GACAGTAGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 549 ccaagatggggaggttataagtcaccccgagcgtgcgtcactactgtactgtcagtgca 608
Db 361 CCAATCGGGAGGTTAAATGATCACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTCA 420
Qy 609 ctctgataagggaagagctgtctactaagcctgagcctgctgtgtgtgtgtgtgtgtgc 668
Db 421 CTTTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Qy 669 cctgcgctgcctgcgaagaaattctcagcactgcgagcagctccctgcggcccgagctcg 728
Db 481 CCGCGGCTGCTGGAGGAATTCCTCAGCAGTGGGCGGAGTTCCTCGGGCCCGACGCTCG 540
Qy 729 cctctgcgaagtgctgtgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 788
Db 541 CTTCTGCGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 599
Qy 789 cctccctgcggcccatctcagaagctgcgcccctctcactcactcactcactcactcact 848
Db 600 CTTCCCTCGGGCCCATCTCAAGGCTG-CCCTTCTCTACACTTCCGACTTCTTCCAGGT 658
Qy 849 tcaactgagggccctgtgtctcccaagctgtcccaagctgcggcctccctcgaagcgt 908
Db 659 TCACGAGAGGGGCGCTGTCTCCCGCAGTGTGCGGAGGCTGCGGCTGCGGCTGCGGAG- 717
Qy 909 ctctgcggcacc 919
Db 718 CTCTGGGCGACC 728

RESULT 5
BI966060/c 609 bp mRNA linear EST 23-Oct-2001
LOCUS 1e72g04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
DEFINITION cDNA 3' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF
APOPTOSIS ;, mRNA sequence.
ACCESSION BI966060
VERSION BI966060.1 GI:16340465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 609)
Melson,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Helmshka,I., Scaerle,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hallier,A., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvilli,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557

```

```

Email: dmelson@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 412.
location/Qualifiers
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
BASE COUNT 142 a 128 c 209 g 129 t 1 others
ORIGIN
Query Match 42.6%; Score 584.4; DB 10; Length 609;
Best Local Similarity 99.7%; Pred. No. 3.2e-102;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 787 accctccctgggcccactcagaagctgcccctcccaactcactcgaactcctcag 846
Db 609 ACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCTCACCATTGTGGACTTTCAG 550
Qy 847 gtcaactgagggccctgtgtctcccaagctgtcccaagctgcggcctccctcgaagc 906
Db 549 GTTCACTAGAGGGGCCCTGTGTCCTCCCGCAGTGTGTCNCAAGGCTGGCGGCTCCG 490
Qy 907 ctctctggcaccggttccctctgtcccaacccctcgaagcgtcttctgtcgaagcgtcc 966
Db 489 CTCTGTGGGACACCGGCTCCCTCTGCCCCACCCCTCAGCGGCTCTTGTCTCCAGAC 430
Qy 967 cctccctctagaagctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1026
Db 429 CCTCCCTTAGAGGCTGCTGGGCTGTTCACGCTGTTCCTTCACGACATTAATGAC 370
Qy 1027 ttcccaacttacttataaactcccccacacgcccactcactcactcactcactcact 1086
Db 369 TTCCACACTTATCTTAAACATCCCCACCGGCGCATCTCCACCTACAGTACGCTCC 310
Qy 1087 cccctaaccttgagggcccccagtgatctgaactcccccctgtgcacagagaccccaaggc 1146
Db 309 CCTACACCTTTAGAGCCGCCAGTATCTGACCTCCCTGGCCACAGACCCCAAGGCG 250
Qy 1147 attgtgtcaactgtactctgtgtggaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1206
Db 249 ATTGTGTTCACTGTACTCTGTGTGGGCAAGATGTGGTCCAGACACCCCACTTAGG 190
Qy 1207 agaagggctggaactgcgtgcgaggaagcagaagactgtggtcctagcgaagagtgctca 1266
Db 189 AGAGGGGCTGAGACTGTGGGCGGAGGAAGCAAGACATGTGGCGGAGGAGGATTTCCA 130
Qy 1267 aatgtgaggggctgagaagaagaagaagctctcctctgtgaaattcctctgtgtatttaa 1326
Db 129 AATGTGAGGGGGGAGAAACAAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTTAA 70

```

OY 1327 aacagataatatttattatattatgtgacaaaatgtgataaattg 1373
|||||
Db 69 AACAGATATTATTATTATTATTGTGACAAATGTGATAATG 23

RESULT 7
AF163779 1027 bp DNA linear GSS 29-AUG-2000
LOCUS AF163779 Human Homo sapiens genomic clone BAC750E14, DNA sequence.
DEFINITION AF163779
ACCESSION AF163779
VERSION AF163779.1 GI:5726439
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1027)
AUTHORS Cousin, P., Billotte, J., Chaubert, P. and Shaw, P. H.
TITLE Physical map of 17p13 and the genes adjacent to p53
JOURNAL Genomics 63 (1), 60-68 (2000)
MEDLINE 20130114
COMMENT Contact: Shaw PH
Experimental Oncology
Institute of Pathology
Rue du Bugnon 25, Lausanne, VD 1011, Switzerland
sub_clone=AB2R Asc-BamHI PSL1180
Class: BAC subclone.
FEATURES
source location/Qualifiers
1..1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="17p"
/clone="BAC750E14"
/clone_11b="Human"
BASE COUNT 207 a 317 c 282 g 208 t 13 others
ORIGIN

Query Match 42.3%; Score 581.4; DB 12; Length 1027;
Best Local Similarity 98.0%; Pred. No. 1.3e-101;
Matches 577; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 785 gaacccctccctgagccacatcgaagctgcccctcctacatcctcgagctctcc 844
|||||
Db 1 GCACCCTCCCTGGCCGCAATCTCAAGCTDYCCCTTCCGACCTTGGACTCTCC 60

OY 845 aggttaactgaggggccctgctcctcccaagctcctccagctgcccctccagc 904
|||||
Db 61 AGGTTACTGAGGGCCCTGGTCTCCCAAGTGTCCAGGCTGCGGCTCCCTCGAC 120

OY 905 agctcctcggaccccgctccctcctgcccacccctcgaagcagctcttgcagagctg 964
|||||
Db 121 AGCTCTGTGGGACCCGGTCCCTCTGYCACACCTCAGCGCTCTTGTCCAGACTG 180

OY 965 cccctcctcctagagctgctgagctgcttcaagcttccatccacataaatacag 1024
|||||
Db 181 CCCCTCCCTCTAGAGGCTGCTGGCTGCTTTCACGTGTTTCCATCCCAATAAATACAG 240

OY 1025 tatcccaactatattttaaactcccccacagccacactcctcaactcacttaactccca 1084
|||||
Db 241 TATTCCACTCTATCTTCAANNNNNNCCACGCCCACTTCCACCTCACTTACCTCCCA 300

OY 1085 atcccgaccccttgagcccccagctgactcgaactcccccctgagcagaccccaag 1144
|||||
Db 301 ATCCGACACCTTTGAGGSCCCAGTGAATCGACATCCGCCCTGGCCACAGACCCCAAG 360

OY 1145 gcaatgcttactgactgctgctgagcaagatgggtccagaaagacccaactcagagc 1204
|||||
Db 361 GCATTGTGTTCACTGTACTGTGGGCAAGATGGGTCCAGAAAGACCCCACTTCAAGCAC 420

OY 1205 taagaaagagctgagcctgagcaggaagcaggaagactgggctccagcagaagatcc 1264
|||||
Db 421 TAAGAGGGGCTGAGCCTGGGGGAGGAAGCAAGAAAGACTGGGCTTAGGGCAGAGATTCC 480

OY 1265 caaatgtaggggagcgaagaacaagaagctcctccctgagagaattccctgtagatttt 1324
|||||
Db 481 CAAATGTAGGGGGCGAGAAACAAGACAACTCCTCCCTTGAGAAATTCCTGTGGAATTTT 540

OY 1325 aaaaagataatatttattattattatgtgacaaaatgtgataaattg 1373
|||||
Db 541 AAAACAGATATTATTATTATTATTATTGTGACAAATGTGATAAATG 589

RESULT 8
BG110063 910 bp mRNA linear EST 30-JAN-2001
LOCUS BG110063
DEFINITION 602279667P1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4367225 5',
mRNA sequence.
ACCESSION BG110063
VERSION BG110063.1 GI:12603569
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC <http://mgi.mgi.nhl.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nhl.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LHAM10019 row: e column: 18
High quality sequence stop: 493.
FEATURES
source location/Qualifiers
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/IMAGE="IMAGE:4367225"
/clone_11b="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: PCMV-SPOrt6; Site:1; NCI;
Site:2; Salt: Cloned unidirectionally; oligo-dr primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 262 a 279 c 225 g 144 t
ORIGIN

Query Match 38.7%; Score 531.6; DB 10; Length 910;
Best Local Similarity 88.6%; Pred. No. 4.6e-92;
Matches 622; Conservative 0; Mismatches 74; Indels 6; Gaps 4;

OY 627 tgcctacccgaagctgagcttgctgtagatgltgctgagccctgagctgagcagga 686
|||||
Db 1 TGCTACCTGAACTGAGCTGTGTGTGATGATGTGTGTGCTGCGCTGCTGAGGA 60

OY 687 atctcaagcaactgagcagatccctcctgagcccccagctcgcgctcctgcaagtgctcg 746
|||||
Db 61 ATTCTCAGCACACGCGGCCAGTTCCCTGGGCCCAAGCTCCGCTGTGCAAGTGTCTGG 120

OY 747 gctgttgcccttgagcagaggtcctcctgagatcagacacccctccctgggcccatt 806
|||||
Db 121 GCTGTGGCCCTGCGGCA-GGTCTCCCTGGGAGATCCGACACCTCCCTGGGCCCATCT 179

OY 807 caaggtgcccccttctcactcacttgagactcttcagagtgtaactagggagccctgt 866
|||||
Db 180 CAAGGCTGCCCTTCTCACCCTTTCGAGACTTTCAGAGTTCACTTGAAGGGGCCCTGGT 239

FEATURES	source
1. .531	Location/Qualifiers
High quality sequence stop: 579.	
plate: LLAM1447 row: 3 column: 23	
found through the I.M.A.G.E. Consortium/LLNL at:	
Clone distribution: MGC clone distribution information can be	
DNA sequencing by: Incyte Genomics, Inc.	
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
CDNA Library Preparation: Life Technologies, Inc.	
Tissue Procurement: Life Technologies, Inc.	
Contact: Robert Strausberg, Ph.D.	
unpublished (1999)	
National Institutes of Health, Mammalian Gene Collection (MGC)	
NIH-MGC http://mgc.ncl.nih.gov/ .	
1 (bases 1 to 531)	
Mammalia; Euteleostomi; Chordata; Vertebrata; Euteleostomi;	
Eumetazoa; Metazoa; Primates; Catarrhini; Hominiidae; Homo.	
ORGANISM	Homo sapiens
SOURCE	human.
KEYWORDS	EST.
VERSION	BI824443.1 GI:15935993
ACCESSION	BI824443
LOCUS	603036693F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179510 5',
DEFINITION	mRNA sequence.
RESULT	9 531 bp mRNA linear EST 04-OCT-2001
Db	656 CAAGGAAGACCTCCCATGAGGATGCGCTGAGATATGACAA 697
Oy	1286 aagacaagctctcccttgagaatccctctgagatttttaa 1327
Db	596 AGGGAAGCAAGAGACATGAGGCTTAGGCAAGACATATCCCAATGACGGGCAAGAAA 655
Oy	1227 cagaagaccaaagaagacg-9ccatgagcagaagttcccaattgga99cga9aac 1285
Db	539 GAGGCAAGATGGGTCCAGAAAGAGC--CATGAGGACGCCAAAGAGGGTGAACCTGAGAGC 595
Oy	1167 tgggaagaagatgggtcccaagaagaccacattcaggcactaagaagggtctgagcctg99g 1226
Db	479 AGTGATCTCGATCCCGCCAGTGGACAGACGCCCGCAGATAGTACATGACTACTGT 538
Oy	1107 cagtgatctcgactccccctggccacagacccccagaagcaltgtgtcaactgtactctg 1166
Db	420 CTCGCCCAAGCG-CCACTCTCCACTCTACTAGTCTCCCAATCCCTGATCCCTTTGAGAGCCCC 478
Oy	1047 ctcccccaccccccaacttccacactctagctcccaatcccttgagcccttgaagcccc 1106
Db	360 GGGCGCTGTTCACGTTTTCATCCCATCAATTAATACATATTCACACTTATCTTACAA 419
Oy	987 gggcctgttcaagtggttttccatccacataaatacagatattccacattatcttaaca 1046
Db	300 CTCTGCCCCACCTCAGCGCGCTTTGGTCCAGACTGCGCCCTCCCTCTAGAGCGCTGGCT 359
Oy	927 ctctgccccaccctcaagcgcgtcttctgttccagacctgccccctctctagagctgct 986
Db	240 CTCGCCAAGTGTGCCAGGCTGCCGGCTCCCTTGAAAGCTCTGTGGACACCGGCTCC 299
Oy	867 ctcccacaagtcggtcccaagcgtctgacggtctccctctgaagcctctctggagaccggtccc 926
Db	867 ctcccacaagtcggtcccaagcgtctgacggtctccctctgaagcctctctggagaccggtccc 926

destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library.

Query Match	Similarity	37.6%	Score 516.8	DB 10	Length 531
Best Local	Similarity	99.4%	Pred. No. 2.9e-89		
Matches	529	Conservative	0	Mismatches	2
				Indels	1
				Gaps	1
QY	661	gtctgtgccttgagctgcgtgcctgaggaattcttagcaactgcgcgcagttccctctggccc	720		
DB	1	gtgtctggcccttcctgcgtgccttgaggaaattctcagacactgcggccagttccctcggccc	60		
QY	721	cagcttcgcctcttgccaagtgctctgggctgtttggccctgycggccaaggtctccctcgag	780		
DB	61	cagcttcgccttcgtccagagttgtctgggctgtttggcccttcggccca-ggtccctccgcgg	119		
QY	781	atcgcgacccctccctctggggcccatctccaaggtctgcccctctcaactacttgagac	840		
DB	120	atccccgacccctccctctggggcccatctccaaggtctgcccctctctcactactttggac	179		
QY	841	ttccaggttactatgaggggcccttggtctctcccaacagtctgccaggtctgcggtccct	900		
DB	180	tttcagagtttactatgaggggcccttggtctctcccgacgtctccagaggtctgcggctccct	239		
QY	901	cgacagctctctggtgcaaccggtctccctctgcaccaectcaagcgcctttgtctcaga	960		
DB	240	cgacagctctctggtgcaaccggtctccctctgcaccaectcaagcgcctttgtctcaga	299		
QY	961	cctgccccctccctctagaaggtgcgtgcctgggcctgttcaagttttccatcccaataa	1020		
DB	300	cctgccccctccctctagaaggtgcgtgcctgggcctgttcaagttttccatcccaataa	359		
QY	1021	acagtatcccaactctatcttacaactccccaaccgcgcacatctccactcaactagtc	1080		
DB	360	acagtatcccaactctatcttacaactccccaaccgcgcacatctccactcaactagtc	419		
QY	1081	cccaatccctgacccctttagagcccccagttgactctgactcccccctgycacaaga	1140		
DB	420	cccaatccctgacccctttagagcccccagttgactctgactcccccctgycacaaga	479		
QY	1141	caggcatctgttctactctgttctgttggcacaagatggtctcagaaga	1192		
DB	480	caggcatctgttctactctgttctgttggcacaagatggtctcagaaga	531		

RESULT	10
LOCUS	Bf577781
DEFINITION	Bf577781 918 bp mRNA linear EST 12-DEC-2006 602029280R1 NCI-CGAP-Cc24 Mus musculus cDNA clone IMAGE:4206595 5'
ACCESSION	mRNA Sequence.
VERSION	Bf577781
KEYWORDS	Bf577781.1 GI:11651493
SOURCE	EST.
ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 918) NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rgs@nih-mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be

MGI:1058389
Seq primer: -40RP from Gibco
High quality sequence stop: 433.
Location/Qualifiers

FEATURES

SOURCE

1. 561
/organism="Mus musculus"
/strain="129 - C57/B6 - FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3155633"
/clone_1ib="NCI-CCAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others

ORIGIN

Query Match 33.1%; Score 454.4; DB 9; Length 561;
Best Local Similarity 88.1%; Pred. No. 2.7e-77;
Matches 494; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

OY 121 aaccgagagcgagagggcgccggggggcgagccgcctctggtccgcgtcg 180
|||||
DB 1 ACCCAGAGCGGAGGGGGGCGGGGGGAGCGGCGACCGCCCTGCGCCCGCTGG 60
OY 181 cgggagcctgggagcctggcgctggcgcccgccctcgctcgccgtgagttg 240
|||||
DB 61 CTGAGCCTGGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 120
OY 241 agcgggacatcgctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
|||||
DB 121 ACCTGGGCAACCTGTCTCCAGGAGCCTTCTCAGAGAGACTGACACAGAGACCGC 180
OY 301 caggagccgctgggaactgatacccgagagagagagagagagatctcgcttccg 360
|||||
DB 181 CCGGAGCCCCCTGAACTGAATCCCAAGACAGAGAAAGCAGAGATGTGTACCTTCTTG 240
OY 361 aaccagactatctggcgcctcgacgaagtgaccctaaagcgagaaacagggctcg 420
|||||
DB 241 GAAACAAGTGTCCGGCTCGAAGAAAGTGTCTCTAAAGCGGAGGCGGCGCTCGCCCA 300
OY 421 ggcagtcgagcccatltaagtlcaaccagaccttgagcaggaacgagcagagt 480
|||||
DB 301 GCTATTGACAGCCCATTTATGAGGTTCCTCGGCGCAGGAGATGACACAGCAGAGT 360
OY 481 gggagcggagacagtgagtgctgggagagagagagagagagagagagagagag 540
|||||
DB 361 GGGAGAGGAGACAGTGTGCTGGGAGAGACCAAAATTAACAGCTCCAGCCCTTGTGCC 420
OY 541 tcaaacacgaatcgagagatttataagtcacccgggctgggctctacacactgact 600
|||||
DB 421 TACGAGCGGACAGATTGGGAATTTACAGTCAATCAGGCGCTGAGCTACTGTA 480
OY 601 caggctgacatttgatgagggagagagctgtctacactgaagctggaactgtctg 660
|||||
DB 481 CAGGTGCACATTGATAGCGGAAAGGCTGTACTGAACTGAGCTGCTGTAAGCGT 540
OY 661 gggcggcctggcgtgctgctgctgctgctgctgctgctgctgctgctgctg 681
|||||
DB 541 GTGCTGGCCCTGCGCTGCTG 561

RESULT 13
LOCUS BG686319 587 bp mRNA linear EST 01-MAY-2001
DEFINITION 602638232P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766071 5',

ACCESSION mRNA sequence.
VERSION BG686319
KEYWORDS BG686319.1 GI:13917716
EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LICM1625 row: p column: 08

High quality sequence stop: 587.
Location/Qualifiers

FEATURES

SOURCE

1. 587
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4766071"
/clone_1ib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 108 a 187 c 205 g 87 t

ORIGIN

Query Match 32.3%; Score 442.8; DB 10; Length 587;
Best Local Similarity 99.3%; Pred. No. 4.6e-75;
Matches 455; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 25 tccgcccggcgcctcccccctcccgatccctcgaggtccggagtgaggcggtgagg 84
|||||
DB 17 TCCGCCCGCGGCTCCCTCCCGGATCCCTCGGGATCCCGGAGTGGGGGCGGTGAGG 76
OY 85 caggagacagcccccccgcccaatggcgcgcctggagagcagaagcgaggggagcg 144
|||||
DB 77 CAGGCAACAGCCCCCGCCCAATGGCGCCGCTCGAGACAGAGGGGAGGGGCGCG 136
OY 145 gggagacggagacagccctctgtgttccgcctgcgccttgagctggcctggc 203
|||||
DB 137 GGGAGACCGGGGACCGCCCTGCTGTGTCGCGCTGGCGCTGGCGCTGGCTGG 196
OY 204 ctgctcggcctcctctgctgagcggtgagtgagttggggagcgagacatcgctc 263
|||||
DB 197 CTGCCCTCGGCTCTCGCTGGCTGGCTGAGTGTGGGAGACCGGCGCATCTGTCG 256
OY 264 ggaagcctggcggagagagctggtgagagagagagagagagagagagagag 323
|||||
DB 257 GGAAGCTGCCAGAGAGAGCTGTGGCAGAGAGGAGCAGGACCCGTCGGAATGAT 316
OY 324 ccag 383
|||||
DB 317 CCAGACAGAGAAAGCAGAGATCTGCGGCTTCTCGAAGCAGATGATGTCGCGC 376
OY 384 aagtgaccctaaagggcggagaaacacgggctcgaagagcagatcgaagccat 443

[illegible]

/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;
Site: 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation: average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT 71 a 186 c 86 g 97 t
ORIGIN

Query Match 31.8%; Score 436.8; DB 10; Length 440;
Best Local Similarity 99.5%; Pred. No. 6.3e-74;
Matches 438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 771 ctccctgagatccgcacccctccctggccatctcaagctgccccctctctaccta 830
DB 1 ctccctgccgagatccgcacccctccctggccatctcaagctgccccctctctaccta 60

QY 831 ctctggactcttccaggttcaactgaaggccctgctctcccaagtcgtcccaagctgc 890
DB 61 ctctggactcttccaggttcaactgaaggccctgctctcccaagtcgtcccaagctgc 120

QY 891 cggctccctcgaagactctctgagcaccggtccctctgcccacacctcagcgctct 950
DB 121 cggctccctcgaagactctctgagcaccggtccctctgcccacacctcagcgctct 180

QY 951 ttgctcagagactgcccctccctctgagagctgctgggctgttcaagtgtttccatc 1010
DB 181 ttgctcagagactgcccctccctctgagagctgctgggctgttcaagtgtttccatc 240

QY 1011 ccacataaatacagla tcccaacttlatcttacaactccccaacgcccactctccacc 1070
DB 241 ccacataaatacagla tcccaacttlatcttacaactccccaacgcccactctccacc 300

QY 1071 tcaactagctccccaatccctgaaccttgaagcccccaagtgaactcgaactccccctggc 1130
DB 301 tcaactagctccccaatccctgaaccttgaagcccccaagtgaactcgaactccccctggc 360

QY 1131 caaagaccccccaaggagatgtgttcaactgtaactctgtggcaagagatgggtccagaagac 1190
DB 361 caaagaccccccaaggagatgtgttcaactgtaactctgtggcaagagatgggtccagaagac 420

QY 1191 cccacttcaggcaactaagag 1210
DB 421 cccacttcaggcaactaagag 440

Search completed: June 22, 2002, 10:49:15
Job time: 18070 sec